



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

3D REKONSTRUKCE OBJEKTŮ Z MULTISKOPICKÝCH SNÍMKŮ

3D RECONSTRUCTION FROM MULTIPLE IMAGES

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Jan Řehoř

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Roman Jakubíček

BRNO 2019

Bakalářská práce

bakalářský studijní obor **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

Student: Jan Řehoř

ID: 195747

Ročník: 3

Akademický rok: 2018/19

NÁZEV TÉMATU:

3D rekonstrukce objektů z multiskopických snímků

POKYNY PRO VYPRACOVÁNÍ:

1) Seznamte se s problematikou 3D rekonstrukce obrazu z více fotografických snímků téže scény z různých pohledů. 2) Prostudujte dostupné rekonstrukční algoritmy a zpracujte literární rešerši na danou problematiku. 3) Navrhněte vhodný akviziční, kalibrační a rekonstrukční přístup pro získání série fotografických snímků zvoleného reálného objektu. 4) Realizujte experiment a implementujte vhodné metody 3D rekonstrukce ze získaných snímků. 5) Otestujte zvolené rekonstrukční algoritmy a vliv akvizičních a rekonstrukčních parametrů na kvalitu rekonstrukce 3D objektů. 6) Diskutujte problematiku případy, dosažené výsledky na reálných objektech a možnosti využití 3D rekonstrukce v medicíně.

DOPORUČENÁ LITERATURA:

[1] ŽÁRA, Jiří, et al. Moderní počítačová grafika. Computer press, 2004.

[2] MOONS, Theo, et al. 3D reconstruction from multiple images part 1: Principles. Foundations and trends® in Computer Graphics and Vision, 2010, 4.4: 287-404.

[3] MOONS, Theo; VERGAUWEN, Maarten; VAN GOOL, Luc. 3D reconstruction from multiple images. 2008.

Termín zadání: 4.2.2019

Termín odevzdání: 24.5.2019

Vedoucí práce: Ing. Roman Jakubiček

Konzultant:

prof. Ing. Ivo Provazník, Ph.D.
předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

ABSTAKT

Tato práce se zabývá 3D rekonstrukcí objektů z multiskopických snímků. Je rozdělena na pět kapitol. První kapitola se týká teoretického seznámení problematiky, tedy, co je 3D rekonstrukce, popis geometrie kamery a rekonstrukční algoritmy. Dále je práce tvořena praktickou částí, ve které jsou popsány jednotlivé kroky realizace 3D rekonstrukce, tedy akvizice snímků, kalibrace a použitých algoritmech. Poslední část práce tvoří diskuze a závěr, ve které jsou diskutovány a shrnuty výsledky pokusu.

KLÍČOVÁ SLOVA

3D rekonstrukce, počítačové vidění, Harris-Stephen algoritmus, SfM, rekonstrukce ze siluet, aktivní stereo rekonstrukce ze strukturovaného osvětlení, Zhengyou Zhang metoda kalibrace, Shape from shading

ABSTRACT

This thesis deals with 3D reconstruction of objects from multiple frames. Thesis is divided into five chapters. The first one is connected with theory, so, it is about 3D reconstruction, about camera geometry description and about reconstruction algorithms. The following are made up by practical parts, in which are described single steps of 3D reconstruction realization, so, it is about frame acquisition, camera calibration and about used algorithms. Last part of the thesis is made up by discussion and conclusion in which the results of the experiment are discussed and summarized.

KEY WORDS

3D reconstruction, Computer vision, Harris-Stephen algorithm, SfM, Shape from silhouettes, Shape from shading Active stereo with structured light, Zhengyou Zhang method of calibration

ŘEHOŘ, Jan. *3D rekonstrukce objektů z multiskopických snímků* [online]. Brno, 2019 [cit. 2019-05-22]. Dostupné z: <https://www.vutbr.cz/studenti/zav-prace/detail/118305>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Roman Jakubíček.

PROHLÁŠENÍ

Prohlašuji, že svoji bakalářskou práci na téma 3D rekonstrukce objektů z multiskopických snímků jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědom následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne

.....

(podpis autora)

PODĚKOVÁNÍ

Rád bych poděkoval vedoucímu bakalářské práce panu Ing. Romanu Jakubíčkov, za pedagogickou a odbornou pomoc a za další cenné rady při zpracování mé bakalářské práce.

V Brně dne

.....

(podpis autora)

OBSAH

ÚVOD	6
1 TEORETICKÝ ROZBOR.....	8
1.1 GEOMETRIE KAMERY	8
1.2 ZPRACOVÁNÍ OBRAZŮ	14
1.3 REKONSTRUKČNÍ ALGORITMY	24
2 NÁVRH.....	29
2.1 METODA „STRUCTURE FROM MOTION“	29
2.2 METODA „SHAPE FROM SILHOUETTES“	30
3 REALIZACE.....	32
3.1 METODA „STRUCTURE FROM MOTION“	32
3.2 METODA „SHAPE FROM SILHOUETTES“	37
4 VÝSLEDKY	40
4.1 METODA „STRUCTURE FROM MOTION“	40
4.2 METODA „SHAPE FROM SILHOUETTES“	47
5 DISKUSE.....	52
ZÁVĚR.....	59

ÚVOD

Tato práce se zabývá 3D rekonstrukcí objektů z multiskopických snímků. Jedná se o metody, které z několika 2D snímků vytvoří 3D obraz (model). Využití těchto metod je rozmanitější, než by se mohlo na první pohled zdát. S touto problematikou se setkal již téměř každý. Za nejviditelnější příklad využití 3D rekonstrukce (z pohledu široké veřejnosti) lze zmínit například odvětví průmyslu volného času. Zejména pak filmový (CGI¹) a herní průmysl, kde aplikace této problematiky urychlilo, dříve, časově náročné modelování uměleckým personálem a umožnilo tak získat, téměř, fotorealistické výstupy. Dále lze zmínit použití ve strojírenství (robotika) anebo v geomorfologii² pro odhad georeliéfu³.

Neméně významná práce s rekonstrukčními algoritmy je i v oblasti bioinformatiky, kde její využití je čím dál tím rozmanitější. Jeho možné využití je například v zachycení pohybů (z angl. motion tracking) a pro kontrolu správného postoje pacienta, které má uplatnění při jeho rehabilitaci, například po vážném zranění [23].

Další možné využití nachází například v protetice. 3D vizualizací zdravé končetiny a následným zrcadlením 3D modelu v software určeném pro 3D modelování⁴ lze získat model chybějící končetiny. Tento model lze následně využít v oblasti 3D tisku respektive biotisku, který je nicméně stále předmětem výzkumu a bývá v klinické praxi uplatněn jen ojediněle. V případě 3D tisku lze s využitím 3D rekonstrukce (a dalšími možnými kroky, které byly zmíněny výše) navrhovat protézy (tiskovým materiálem bývá titan).

Zmíněný proces, biotisk, je podobný “klasickému“ 3D tisku, nicméně využívá biologický tiskový materiál (dále jen bioinkoust). Lze jej využít pro tisk tkání a orgánů. Bioinkoustem může být například kolagen, kyselina hyaluronová, PVA (polyvinilalkohol⁵), kmenové buňky a jiné materiály. Tisk lze následně provést několika způsoby, například využitím technologie FFF (Fused Filament Fabrication) anebo metodou, která je založena na injekčním vytlačování bioinkoustu. V případě prvního zmíněného je řeč o technologii založené na extrudaci roztavené tiskové struny (filamentu) z tiskové hlavy. V případě biotisku lze využít více tiskových hlav, kdy jedna tisková hlava extruduje stavební materiál (v případě biotisku - bioinkoust) a druhá materiál podpůrný. U druhé metody, injekční, je proces odlišný. Tiskovou hlavu zde totiž tvoří injekce, obsahující bioinkoust (například již zmíněný PVA), který je vytlačován (například) na miskou, umístěnou na platformě. Vytlačený bioinkoust bývá ještě zesílen chemickou anebo UV cestou pro získání lepších mechanických vlastností [22]. Možné uplatnění kombinace biotisku a 3D rekonstrukce lze zmínit například v plastické chirurgii, kde lze biotiskem pacientovi vytisknout chybějící končetinu, například ucho.

¹ CGI - Computer-generated imagery - počítačem generovaná grafika a efekty ve filmech

² Geomorfologie - nauka o zemském povrchu a jeho vývoji

³ Georeliéf - Vzhled zemského povrchu

⁴ Blender, Cinema 4D, 3ds Max, SolidWorks, SketchUp...

⁵ Syntetický, ve vodě rozpustný, polymer. Je netoxický a biologicky odbouratelný [24]

Z výše zmíněného vyplývá, že problematika využití 3D rekonstrukce je velmi rozmanitá, neméně výrazným ukazatelem jeho uplatnění je i v diagnostice. Jelikož se ale zejména jedná o přístupy, zpracovávající DICOM soubory, nebudou předmětem zájmu této práce.

V bakalářské práci je věnována pozornost teoretickému i praktickému seznámení do problematiky 3D rekonstrukce z 2D snímků v běžně využívaných formátech (JPG, PNG apod.). Teoretická část (kapitola 1) se věnuje popisu geometrie kamery a rekonstrukčním algoritmům, které s 3D rekonstrukcí souvisejí. V praktické části (kapitoly 2 - 3) je věnováno výběru a návrhu rekonstrukčních přístupů a realizaci experimentu, ve kterém jsou navrhnuté přístupy otestovány.

Poslední část (kapitoly 4 - 5) se věnuje diskuzi získaných výsledků a jejich porovnání s výstupem z komerčního programu.

1 TEORETICKÝ ROZBOR

Cílem této kapitoly je teoretické seznámení s problematikou 3D rekonstrukce. Je rozdělena na tři části. První část pojednává o geometrii kamery, druhá podkapitola se zaměřuje na zpracování obrazů a třetí podkapitola se věnuje již samostatným rekonstrukčním algoritmům.

1.1 GEOMETRIE KAMERY

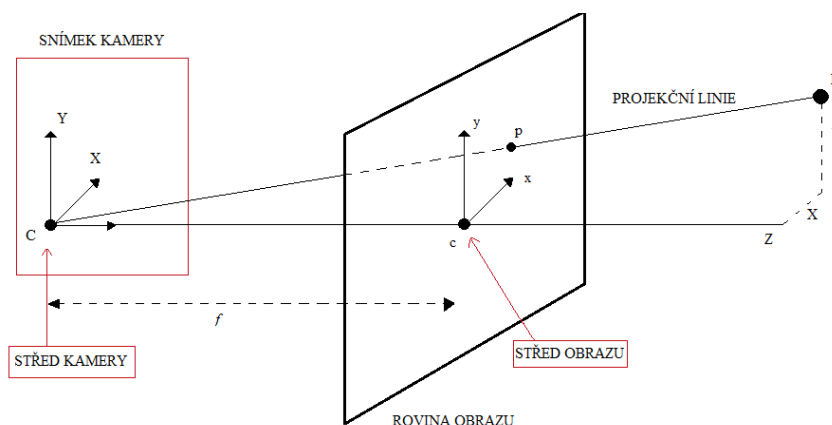
Během realizace a aplikací rekonstrukčních algoritmů pro 3D vizualizaci, se pracuje s geometrickými vztahy mezi sadou snímků a 3D scénou.

V počítačové grafice se trojrozměrné objekty zobrazují na dvojrozměrných zobrazovacích zařízeních (fotoaparát, videokamera). Tato transformace, tedy převod z 3D objektu na 2D reprezentaci se nazývá promítání, neboli projekce [1].

Tato kapitola bude věnována geometrii kamery. V první části budou zmíněny projekce kamery, tedy perspektivní (středová) a ortografická (paralelní) projekce. Také budou zmíněny typy zkreslení, vnitřní a vnější parametry kamery a jejich odhad (kalibrace).

1.1.1 PERSPEKTIVNÍ PROJEKCE

Pro tento typ promítání (a stejně tak i u jiného typu) je charakteristická deformace objektů z 3D prostoru do dvourozměrného výstupu. Výsledek je následně takový, že platí nepřímá úměrnost mezi vzdáleností objektů od kamery a velikostí těchto elementů ve výstupním obrazu. Bývá používán, jako standardní model promítání kamery. Na Obr. 1.1 je ukázka perspektivní projekce na dírkovém modelu kamery⁶.



Obr. 1.1: Perspektivní projekce na dírkovém modelu kamery

⁶ Dírková kamera (z angl. pinhole camera) je fotografické zařízení tvořeno světlotěsnou skřínkou, kde je optický systém (jak bývá u současných fotoaparátů) nahrazen velmi úzkým otvorem. Světlo procházející otvorem dopadá na fotografický film. Princip dírkové kamery byl znám už ve starověku.

Střed promítání (optický střed kamery) zde reprezentuje bod C, bod c je středem obrazu (průmětny), který je rovněž průsečíkem hlavní osy. Vzdálenost mezi středem promítání a středem průmětny se nazývá ohnisková vzdálenost (např.: vzdálenost mezi čočkou a snímačem). Značí se f .

Trojrozměrný bod $P = (X, Y, Z)^T$, se promítá do dvojrozměrného bodu $p = (x, y)^T$ dle vztahu

$$x = f \frac{X}{Z}, y = f \frac{Y}{Z} \quad (1.1)$$

Aby bylo možné vykonat veškeré grafické operace⁷, je potřeba, aby pro zpracování grafickými kartami byly reprezentované, jako násobení matic. Aby bylo možné použití maticových operací, je třeba přidat další, homogenní, souřadnice w , která se nazývá váha bodu. Ta je většinou rovna 1.

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.2)$$

Ze vztahů 1.2 a 1.1 tedy plyne, že

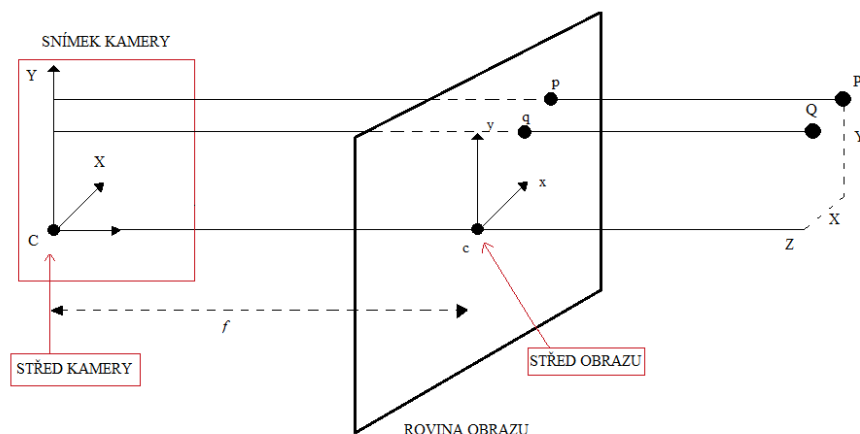
$$\begin{aligned} x &= \frac{x_h}{w} = f \frac{X}{Z} \\ y &= \frac{y_h}{w} = f \frac{Y}{Z} \end{aligned} \quad (1.3)$$

Dle obrázku 1.1 by paprsek, procházející souřadnicí kamery by pokračoval dále.

1.1.2 ORTOGRAFICKÁ PROJEKCE

Základním rozdílem ortografické projekce od středové je ten, že objekt nehledě na vzdálenost od projekčního systému, nemění svou velikost. Všechny projekční linie jsou paralelní k hlavní ose – viz Obr.1.2.

⁷ Elementární grafické operace jsou translace (posunutí), rotace (otočení), „scaling“ (změna měřítka) a „shear“ (zkosení). Pro translaci platí vztah $P' = P + T$, kde P je matice reprezentující výchozí pozici bodu, P' reprezentuje novou pozici. Použitím homogenních souřadnic lze translaci, tedy součet P a T vyjádřit násobením matic [1].



Obr. 1.2: Ortografická projekce kamery

Trojrozměrné body $P = (X_p, Y_p, Z_p)^T$ a $Q = (X_q, Y_q, Z_q)^T$ se promítají do dvojrozměrných bodů $p = (x_p, y_p)^T$ a $q = (x_q, y_q)^T$ dle obecného vztahu

$$\begin{aligned} x &= X \\ y &= Y \end{aligned} \quad (1.4)$$

Pro rychlé zpracování grafickými kartami se však doporučuje použití homogenních souřadnic, ty lze maticově vyjádřit následovně.

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.5)$$

Ze vztahů 1.4 a 1.5 tedy plyne, že

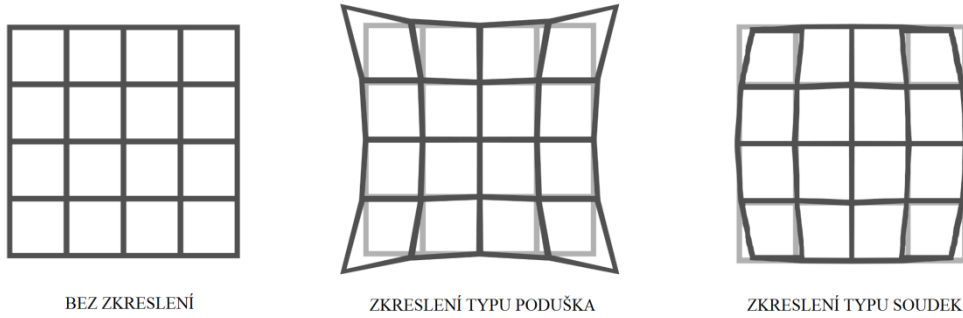
$$\begin{aligned} x &= \frac{x_h}{w} = X \\ y &= \frac{y_h}{w} = Y \end{aligned} \quad (1.6)$$

1.1.3 RADIÁLNÍ ZKRESLENÍ SNÍMKŮ

V případě, kdy je vyžadován přesný obrazový výstup, se musí počítat s odchylkami, které mohou nastat během projekce. Tyto odchylky se projevují tak, že úhel vystupujícího paprsku není shodný s úhlem paprsku vstupujícího.

V případě 3D rekonstrukce tento fakt rozhoduje o kvalitě výstupu. V geometrické optice se tento jev nazývá distorze (zkreslení). Tento jev bývá nejčastěji způsoben nedokonalostí čoček u objektivů, případně volbou nízké ohniskové vzdálenosti při expozici snímku. Zkreslení může být perspektivní, „skew“ (mimoběžné), nejvýraznější jsou nicméně radiální distorze. Ty se projevují přemístěním bodů (souřadnic) od středu obrazu v závislosti

na jejich vzdálenosti. Přemístění může být buď od (poduška) anebo ke středu (soudek) pozorovaného obrazu v závislosti na radiální vzdálenosti⁸ – viz obr. 1.3.



Obr. 1.3: Radiální zkreslení

Radiální distorze je modelována následujícím vztahem [2][12]

$$\begin{aligned} x &= x_c(1 + k_1 r_c^2 + k_2 r_c^4) \\ y &= y_c(1 + k_1 r_c^2 + k_2 r_c^4) \end{aligned} \quad (1.7)$$

přičemž, $r_c^2 = x_c^2 + y_c^2$,

kde $(x_c, y_c)^T$ jsou souřadnice nezkreslených bodů v závislosti na vzdálenosti od středu obrazu. Zkreslené souřadnice bodu jsou reprezentovány $(x, y)^T$. Znaky k_1 a k_2 jsou radiálně distorzní parametry. Většinou se odhadují, zpravidla ale k_1 popisuje 90% distorzi.

1.1.4 PARAMETRY KAMERY

Pro realizaci 3D rekonstrukce je důležité znát informace o kameře, tedy, jak je daleko od snímaného objektu, jakou má ohniskovou vzdálenost a mnohé další, díky kterým lze odhadnout 3D strukturu z pixelových souřadnic bodů z 2D snímků. Tyto vlastnosti se souhrnně nazývají parametry kamery. Parametry kamery lze rozdělit na dvě skupiny, vnitřní a vnější.

VNĚJŠÍ PARAMETRY

Vnější parametry popisují lokaci a rotaci kamery v prostoru. Jsou definovány translačním vektorem t a rotační maticí R ,

$$\begin{aligned} t &= [t_1, t_2, t_3]^T, \\ R &= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \end{aligned} \quad (1.8)$$

vnější parametry kamery jsou pak popsány transformační maticí [2][29]

⁸ Radiální vzdálenost - vzdálenost od středu směrem k okraji snímku.

$$M_{ex} = [Rt]$$

$$M_{ex} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \quad (1.9)$$

VNITŘNÍ PARAMETRY

Popisují geometrické, digitální a optické vlastnosti kamery. Vnitřní parametry kamery jsou popsány maticí [5]

$$M_{in} = \begin{bmatrix} \beta f & \gamma f & o_x \\ 0 & f & o_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1.10)$$

, kde f je ohnisková vzdálenost (px), β je tzv. aspect ratio – neboli poměr stran pixelů, popisuje vztah mezi jeho šířkou a výškou. Dále se v matici nachází symbol γ , který popisuje zešíkmení pixelů. Znaky o_x a o_y jsou tzv. principal points, neboli souřadnice hlavního bodu [30].

1.1.5 KALIBRACE KAMERY

Kalibrace kamery je stěžejním krokem 3D vizualizace, jelikož díky ní lze vypočítat vnější a vnitřní parametry kamery, které jsou nezbytné pro další práci. Jinými slovy, je odhadnuta nejen prostorová informace, tedy poloha kamery v prostoru, ale jsou rovněž odhadnuty informace, které popisují zařízení samotné. Jedná se například o ohniskovou vzdálenost, velikost pixelů anebo také o distorzní koeficienty, poskytující informaci, kterou lze využít pro korekci zkreslení. Tuto techniku lze klasifikovat do dvou kategorií, na fotogrametrickou kalibraci a auto-kalibraci (z angl. self-calibration).

Pro první zmíněnou kalibraci je typické, že využívá kalibračního objektu, kde je známa prostorová informace. Volí se nejčastěji šachovnice⁹, jelikož je pro zpracování, díky své pravidelnosti a vzhledu, efektivní. Někdy se volí levý dolní roh, jako bod, reprezentující referenční souřadnice prostředí. Některé metody využívají těchto objektů více, nejčastěji dva až tři, které bývají na sebe kolmé.

Druhá kategorie, tedy auto-kalibrace, je schopna provést kalibraci kamery bez potřeby kalibračního objektu. Pro kalibraci využívá informace z obrazů (příznaky) [3]. Auto-kalibrace je využívána u většiny komerčních programů¹⁰, jelikož tato technika poskytuje konečnému uživateli pohodlnější práci.

I přesto, že programy využívající auto-kalibraci, nevyžadují kalibrační objekt, mohou nastat případy, kdy bez kalibračního objektu není 3D vizualizace možná. Může se jednat o objekty, které mají jednobarevnou texturu. Tím, na rozdíl od objektů, které mají pestré textury, komplikují detekci příznaků obrazu.

⁹ Šachovnice je nejvíce využívána, nicméně lze využít jiný kalibrační objekt, většinou se ale jedná o rovnoměrně rozmístěné černé čtverce na bílém pozadí

¹⁰ Agisoft Photoscan, Autodesk® Memento™, 3DF Zephyr

Je několik přístupů pro výpočet parametrů kamery. Lze zmínit například metodu Faugeras-Toscani, či metodu Zhang. Obě metody využívají algoritmy pro detekci rohů. Zhang kalibrace vyžaduje nejméně 2 snímky šachovnice z různých pohledů, z pravidla se však používá 10 až 30 snímků. Metoda Faugeras-Toscani používá tzv. „non-plannar patter“, jedná se o 3D kalibrační objekt. Většinou tvořen dvěma až třemi šachovnicemi postavenými kolmo k sobě. Metoda Zhang pracuje s 2D kalibračním objektem („plannar pattern“), ta může mít podobu jedné šachovnice vytisknuté na papíře [4]. Výhodou metody Zhang je jednoduchost, díky využití 2D kalibračního objektu, zatímco metoda Faugeras-Toscani vyžaduje složitější kalibrační objekt, který je na realizaci náročnější. Rozdíl mezi non-plannar a plannar pattern je ilustračně znázorněn na obrázku 1.4.



Obr. 1.4: Kalibrace kamery. Využití non-plannar pattern (vlevo) a využití plannar pattern (vpravo)

Jak již bylo zmíněno na začátku této podkapitoly, kalibrace slouží ke stanovení vnitřních a vnějších parametrů kamery. Je-li 3D bod v souřadnicích definován jako $X = [X, Y, Z, 1]^T$, změna měřítka symbolem s , matice vnitřních parametrů, jako M_{in} , vnějších parametrů M_{ex} a 2D v obrazu, jako $x = [x, y, 1]^T$, lze geometrickou kalibraci kamery popsat matematicky následovně

$$sx = PX \quad (1.11)$$

, přičemž

$$P = M_{in}M_{ex} \quad (1.12)$$

, kde P je kalibrační matice kamery.

Vzorec (2.11) lze rozepsat podrobněji

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} \beta f & \gamma f & o_x \\ 0 & f & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1.13)$$

1.2 ZPRACOVÁNÍ OBRAZŮ

V případě 3D rekonstrukce je mnohdy nezbytně nutné mít alespoň elementární znalosti týkající se zpracování obrazů. Jelikož je tato problematika velmi rozsáhlá, budou zmíněny zejména ty operace, které byly využity v rámci praktické části bakalářské práce.

1.2.1 DETEKCE OBRAZOVÝCH PŘÍZNAKŮ

Detekce obrazových příznaků spadá do prvních kroků, které se vykonají před řešením složitějších úloh (analýza obrazu, využití v počítačovém vidění) [2][8]. Lze je chápat, jako specifické dílky textury poskytující informace pro následující číslíkové zpracování [6].

Umožňují počítači (případně robotovi) interpretovat informace ze získaných snímků. Volba vhodného přístupu má tedy zásadní vliv na kvalitu výstupu. Uplatnění je velmi rozmanité, například při identifikaci otisků prstů, u kamer, určování metrických vlastností objektů, pro tvorbu panoramat ze dvou a více fotek, pro práci s medicínskými obrazy anebo jako základ pro segmentační přístupy [15]. Jako hlavní detekční přístupy jsou například detekce rohů a hran.

Díky detekci rohů lze v 3D rekonstrukci vyhledat vzájemnosti mezi vícero snímky, které jsou velmi důležité pro následné operace, díky kterým lze zrealizovat požadovaný výstup, například ve formě trojrozměrného mračna bodů („point cloud“).

Hrany lze využít, jako základ pro segmentační přístupy [8], jejichž uplatnění může být vhodné pro vytvoření siluet ze vstupních snímků pro „shape-from-silhouette“ (rekonstrukce 3D struktury ze siluet).

HRANOVÁ DETEKCE

Jedná se o metody, sloužící ke zvýraznění obrazových elementů, u nichž je dostatečně rychlá změna jasu. Čím je rychlost změny vyšší, tím je vyšší modul gradientu a tak rovněž i vyšší strmost. V případě digitálních obrazů hrana nemusí být dána pouze oddělením objektů, může být rovněž způsobena změnou barvy povrchu, tvarem, stínováním a jinými faktory [8]. Na základě jasových profilů lze hrany rozdělit na skokové, střechové, linie a zašuměné. Hrana je dána velikostí (1.14) a směrem (1.15) [9],

$$|\nabla_g| = \sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2} \quad (1.14)$$

$$\varphi = \arctg\left(\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}\right) \quad (1.15)$$

přičemž velikost hrany je shodná s velikostí gradientu obrazové funkce [9].

Jelikož se jedná o gradientní metodu, čili, hledá místo s nejvyšší rychlostí změny pohybu, využívají tyto metody derivací, které lze rozdělit, buď na základě vyhledání lokálního

maxima první derivací obrazové funkce anebo průchodem nulou druhou derivací. V diskrétních systémech jsou derivace nahrazeny diferencemi pomocí lokálních operátorů¹¹.

Pro aproximaci první derivace lze využít masky, jako například operátor Sobelův (1.16), Kirschův či operátor Previtové, které bývají označovány také, jako tzv. „kompasní detektory“, protože detekují hrany v určitém směru, proto jsou operátory ve dvou variantách, pro horizontální (h_1) a vertikální směr (h_2). Filtrace tak probíhá dvakrát v těchto směrech [8].

$$h_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, h_2 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (1.16)$$

Pro aproximaci druhé derivace lze využít například operátor Laplacián (1.17) a Laplacián Gaussianu (1.18).

$$L_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, L_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (1.17)$$

$$LoG = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \end{bmatrix} \quad (1.18)$$

Oproti první skupině jsou operátory aproximující druhou derivaci schopné detekovat hrany ve více směrech.

Po získání hodnoty gradientu ze všech pixelů je následně nutné použít prahování pro získání hranové reprezentace. Vznikne tak binární obraz, kde číslo jedna vyjadřuje přítomnost, nula nepřítomnost hrany [8][15]. Ukázka detekce hran pomocí operátorů aproximující druhou derivaci je na obrázku 1.5.

Jak již bylo v úvodu zmíněno, hrany nejsou způsobeny pouze fyzickým oddělením objektů ve scéně. Mohou tedy být způsobeny tvarem povrchu objektu, respektive jeho detaily, ty mají vysokou frekvenci, ve které se nachází i šum. Ten může být při detekci nesprávně označován, jako hrana. Lze proto použít pokročilejší přístup, tzv. Cannyho hranový detektor, který je oproti zmíněným konvolučním přístupům méně citlivý vůči šumu [8].

¹¹ Operátory lze rozdělit na bodové, lokální a globální. Bodové využívají pro výpočet výstupního pixelu odpovídající pixel u obrazu vstupního. Lokální operátory využívají lokální okolí bodu vstupního. Použitím globálních operátorů je výstupní pixel ovlivněn celým vstupním obrazem. Využívají se pro zpracování obrazu [8]

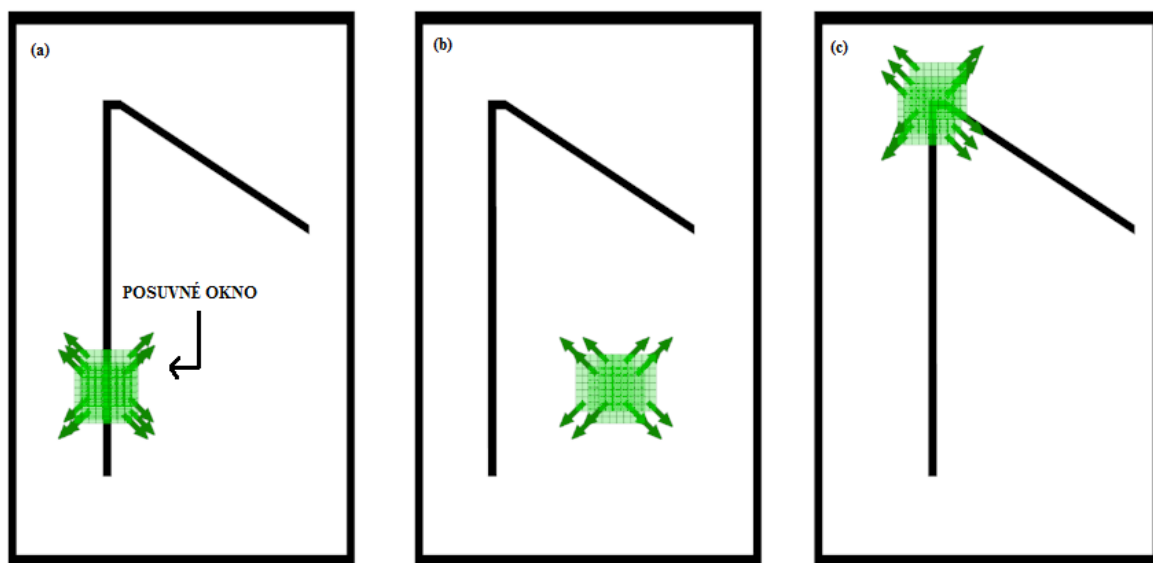


Obr. 1.5: Původní obrázek (vpravo), ukázka hranové detekce aproximující druhou derivaci pomocí lokálního operátoru Laplacián (vpravo)

ROHOVÁ DETEKCE

Při vyhledání vzájemností u dvou a více snímků je potřeba nalézt taková místa, která se co nejméně podobají svému okolí [9]. Pokud se problematika hledání takového bodu ilustrovala pomocí pohyblivého okna, hledající tento bod (případně malou oblast), situace by vypadala následovně (Obr.1.6):

V případě oblasti roviny není téměř možné nalézt body, které by byly nejen specifické, ale rovněž by se obtížně hledala vzájemnost tohoto bodu u snímků jiných. V případě hran je situace lepší, rozdílnou podobu bodu od okolí lze rozeznat ve směru gradientu hrany, ve směru hrany však již nikoliv. Jedinými takto specifickými body jsou rohy.



Obr. 1.6: Ilustrace řešení problematiky vyhledávání vhodného významného bodu pomocí posuvného okna, (a) vyhledávání významného bodu podél hrany, (b) vyhledávání významného bodu na rovině, (c) vyhledávání významného bodu v okolí hrany. Adaptováno z [2]

Lze využít mnoho cest k detekci, jako například Moravcův algoritmus [9][15][17]. **Moravcův algoritmus** (uváděn také, jako operátor) určuje významnost bodu na základě odlišnosti malé části obrazu od jeho okolí pomocí čtvercového okna, které se pohybuje pixel po pixelu [15]. Matematicky je vyjádřen následovně [18]:

$$E(u, v) = \sum_{x,y} w(x, y) [I(x + u, y + v) - I(x, y)]^2 \quad (1.19)$$

, kde $w(x, y)$ je skoková funkce s možnými skoky (x, y) : $(1, 0)$, $(1, 1)$, $(0, 1)$, $(-1, 1)$, $(-1, 0)$, $(-1, -1)$, $(0, -1)$, $(1, -1)$. Vstupem je šedotónový obraz.

Dalším možným přístupem je tzv. **Harris-Stephensonův detektor**, který oproti Moravcovému algoritmu redukoval nedostatky, kterými Moravcův algoritmus disponoval. Je dokonce schopen detekovat hranu. Funkce okna byla skoková. Toto okno bylo citlivé na šum, jelikož se mohlo stát, že šum, který je svojí povahou odlišný od okolí, byl mylně chápán, jako roh. Proto je skoková funkce okna nahrazena Gaussovou funkcí. Harris-Stephensonův detektor je založen na auto-korelační funkci. Je-li definován posun (u, v) je bod auto-korelační funkce definován, jako [17][18]:

$$c(u, v) = \sum_{x,y} w(x, y) [I(x_i, y_i) + I_x(x_i + \Delta x, y_i + \Delta y)]^2 \quad (1.20)$$

Pomocí Tylorovy řady a aritmetických operací je vzorec 1.20 upraven na:

$$c(u, v) \approx [u \ v] C \begin{bmatrix} u \\ v \end{bmatrix} \quad (1.21)$$

, kde C (vzorec 1.22) je auto-korelační matice, která zaznamenává intenzitu struktury z blízkého okolí.

$$C = \sum_{x,y} w(x, y) \begin{bmatrix} I_x I_x & I_x I_y \\ I_x I_y & I_y I_y \end{bmatrix} \quad (1.22)$$

Zde jsou I_x a I_y derivace obrazu, těch lze dosáhnou aproximací pomocí konvolučních masek, jak již bylo zmíněno v hranové detekci.

Nakonec se dle vztahu 1.23 lze následně vypočítat odezvu detektoru O :

$$O = \det(C) - k(tr(C))^2 \quad (1.23)$$

$$\det(C) = \alpha_1 \cdot \alpha_2$$

$$tr(C) = \alpha_1 + \alpha_2$$

, kde k , je konstanta, α_1 a α_2 jsou vlastní čísla (z angl. eigenvalues), $\det(C)$ vyjadřuje determinant auto-korelační matice a $tr(C)$ je tzv. stopa (z angl. trace). Jedná se o součet všech prvků na hlavní diagonále matice.

Na základě odezvy, O , lze stanovit následující závěry: je-li rovno nízkým hodnotám, téměř rovným nule, způsobené nízkými hodnotami α_1 a α_2 , tak se jedná o rovinu. Pokud je odezva v záporných hodnotách, kdy $\alpha_1 \gg \alpha_2$, tak se pak následně jedná o hranu. Nakonec, jsou-li O , α_1 a α_2 rovny vysokým kladným hodnotám, tak se pak jedná o roh [15][17].

Dalším přístupem je Smith Bradyho algoritmus, nazývaný také, jako **SUSAN** (Smallest Univalued Segment Assimilating Nucleus), tato metoda počítá počet pixelů, jejichž hodnota parametru (jasu), je podobná s centrální hodnotou masky, kterou reprezentuje zkoumaný bod (nucleus). Oblast masky, ve které se porovnávají pixely s centrem masky, je označen, jako USAN (Univalued Segment Assimilating Nucleus). Ta obsahuje důležité informace v okolí zkoumaného bodu. Je-li hodnota USAN nízká (nízký počet pixelů s podobnými parametry, jako nucleus), tak se pak jedná o roh. Na Obr. 1.7 je ilustrativní ukázka rohové detekce.



Obr. 1.7: Rohová detekce pomocí Harris–Stephensova algoritmu s 50-ti detekovanými rohy

Harris–Stephensovnův algoritmus je poněkud citlivý na změnu velikosti a nasvětlení obrazu. Jinými slovy, není schopen po přiblížení obrazu nalézt tentýž roh. Je proto třeba tyto body popsat invariantní informací, nezávislou na osvětlení, přiblížení i například na úhlu pohledu. Důležitost nastává zejména při extrakci příznaků ze snímků zachycující stejnou scénu. Tento problém řeší tzv. descriptor, které detekované body charakterizují informací, nehledě na osvětlení, či na jiných faktorech. Díky této skutečnosti je bod v prvním snímku stejně chápán i ve snímku druhém. Jsou jimi například **SIFT** (Scale Invariant Feature Transform), extrahující informaci téhož příznaku nehledě na zvětšení/zmenšení obrazu anebo **SURF** (Speed-Up Robustness Features), který kromě toho, že je rychlejší, jako SIFT, zvládne extrahovat stejnou informaci z téhož bodu nehledě na více transformací [2][6].

Nejen v 3D rekonstrukci, ale i v jiných problematikách, je důležité nalézt mezi dvěma a více snímky korespondující body. Přístupů je mnoho, jedním z nich je metoda **RANSAC**.

Vstupem je sada sledovaných dat, parametrizovaný model a důvěryhodné parametry. Algoritmus probíhá tak, že se vybere náhodná podskupina ze sady dat, ta se z počátku považuje, jako vhodný model. Následně se iterativně testují všechny podskupiny dat s modelem, kde se sleduje vzdálenost mezi sebou. Hodnocení je vymezeno prahem (u obrazů se může jednat o vymezení vzdálenosti daného bodu v pixelech), který je předem stanoven. Nejlepší výsledek u dané podskupiny je považován, jako nový model, který nahradí model

předchozí. Vyhodnocení je určeno dle předem vytvořeného prahu. Počet iterací je předem stanoven. V případě obrazů lze u dvou a více snímků pomocí RANSAC nalézt z významných bodů pouze korespondující body, přičemž ostatní jsou zamítnuty [9]. Je využitý na výpočet korespondencí v epipolární geometrii (kap. 1.2.3).

1.2.2 SEGMENTACE OBRAZU

Segmentace obrazu patří k základním krokům analýzy obrazu. Jedná se o proces, který rozděluje obraz na více částí, případně vyhledává konkrétní oblast, označeny mohou být pouze hranice segmentů. V tomto případě může být výstupní obraz binární, kde jedna vyjadřuje hledaný segment a nula okolí [8]. Využití je široké, například v medicíně (zpracování obrazů pro diagnostické účely), v umělé inteligenci pro rozeznání objektu ve scéně a v neposlední řadě v 3D rekonstrukci.

Zde má nejvýraznější uplatnění pro tvorbu siluet, které využívá algoritmus shape-form silhouette (rekonstrukce struktury ze siluet). Lze využít mnoho přístupů, zde budou zmíněny pouze některé z nich.

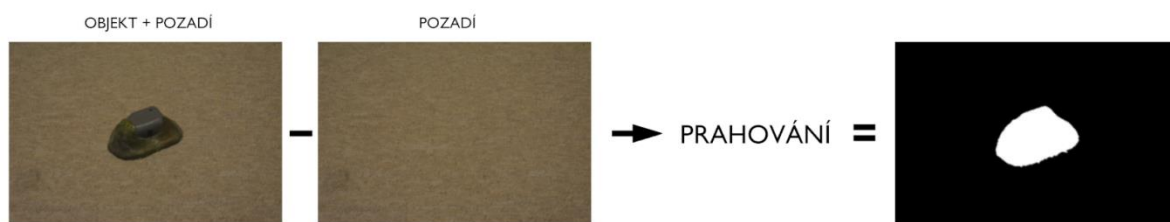
PRAHOVÁNÍ

Jedná se o segmentační metodu založenou na transformaci vstupní g na výstupní obraz f pomocí prahu T [9]. Je-li použitý jen jeden práh, výstupní obraz bude binární (0 – pokud hodnota jasu $< T$; 1: hodnota jasu $> T$) jsou-li použité prahy dva (tzv. dojité prahování), může být výstupní obraz šedotonový. Metoda je citlivá na nasvětlování objektu, kdy například, je-li objekt nasvětlený ze strany a pozorován čelně, může být segmentovaná pouze nasvětlená část objektu [8].

ODEČTENÍ POZADÍ (ZACHOVÁNÍ POPŘEDÍ)

Jsou – li k dispozici dva snímky, kde jeden snímek obsahuje hledaný objekt s pozadím a druhý snímek pouze samotné pozadí (je-li známé i pozadí samotné), lze tento objekt jednoduše vysegmentovat odečtením těchto snímků s případným následným prahováním [20].

Ačkoliv je tato metoda realizačně snadná, má své jisté nevýhody. Metoda totiž u obou snímků vyžaduje stejnou expozici, stejné zaostření a nejlépe, aby byly nasnímány ze stejného úhlu. V jiném případě může dojít k nesprávné segmentaci. Tyto podmínky nelze v terénu snadno zajistit. Rovněž metoda citlivá na nasvětlení objektu, kdy při nesprávné orientaci světla může být segmentován i stín objektu zájmu, což mnohdy nemusí být žádoucí. Rovněž odečtením pozadí není u segmentovaného objektu zachována textura. Ukázka použití metody segmentace odečtením pozadí je na Obr. 1.8.



Obr. 1.8: Metoda segmentace pomocí zachování popředí. Na vstupu jsou dva snímky, kdy jeden snímek obsahuje v popředí objekt zájmu, druhý nikoliv. Jejich odečtením zůstane zachovaný pouze objekt v popředí, pomocí prahování a případné filtrace je vytvořený

METODA NARŮSTÁNÍ OBLASTÍ

Jedná se o segmentační metodu, jejímž cílem je rozčlenění objektu, z hlediska parametru, do souvislých homogenních oblastí [9]. Metoda je realizována tak, že na počátku se stanoví iniciační bod (také nazýván, jako semínko), kterým se kontrolují okolní pixely. Na základě kritéria se následně rozhodne, zdali budou anebo nebudou přidány do vznikajícího segmentu.

Kritéria lze rozdělit na statická a dynamická. Statické kritérium je založené na porovnání absolutní hodnoty rozdílu parametru semínka (jasu) a aktuálně testovaným pixelem s prahem. Po splnění podmínky je pixel přidán do vznikajícího segmentu a v dalším cyklu algoritmu slouží, jako nový výchozí bod. Dynamické kritérium rovněž používá práh, nicméně zde jsou parametry aktuálně testovaného pixelu porovnávány buď s parametry pixelu předchozího anebo se střední hodnotou parametru pixelů v již segmentované oblasti. Tato metoda je méně citlivá na nasvětlení objektu [8].

CHROMATICKÉ KLÍČOVÁNÍ

Další možnou cestou segmentace objektu v obraze je tzv. chromatické klíčování, z angl. chroma keying či greenscreening. Jedná se o postprodukční techniku pro vrstvení dvou snímků/videosekvencí na základě vyjmutí (zprůhlednění pomocí alfa-kanálu¹²) pozadí z jiného snímku. Jiným snímkem je zde myšlena výchozí fotografie/videosekvence, která byla natočena (respektive nafocena) na jednobarevném pozadí. Volí se většinou zelené plátno (proto název „greenscreen“), nicméně barva pozadí může být jakákoliv.

Principem této techniky je zprůhlednění pixelu obsahující stejnou anebo podobnou barvu. Je proto třeba brát na vědomí, aby objekt obrazu, který chceme ve fotografii/videosekvenci zachovat, neměl na sobě tutéž barvu, jako pozadí.

Využití je velmi široké, nejvýraznější je však v zábavním průmyslu, zejména v kinematografii. Objekt zájmu (v tomto případě herec) stojí před monochromatickým pozadím, kde je natáčen. Ze získané videosekvence se odstraní pozadí (bude zprůhledněno) a zachová objekt v popředí (vyklíčuje). Upravený snímek může být následně vložen do 3D

¹² Čtvrtý činitel barevného modelu RGBA – Red, Green, Blue, Alfa. Jedná se o tzv. koeficient průhlednosti. Vyjadřuje, z kolika alfa procent je daný pixel pokrytý neprůhlednou barvou RGB [9]. RGBA využívá například formát PNG.

scény v 3D software anebo navrstven s jiným snímkem (nové pozadí). Zjednodušený příklad barevného klíčování je ilustrován na Obr. 1.9.



Obr. 1.9: Chromatické klíčování, ve vstupním snímku je veškerá zelená vyjmuta, obraz má následné průhledné pozadí¹³, následným vrstvením s novým snímkem vznikne konečný obraz

Stejně tak, jako jiné segmentační metody, má chromatické klíčování ve 3D rekonstrukci své uplatnění. Lze jím vytvořit siluety pro shape-from-silhouettes algoritmus. Postup je identický, jako na obr. 3.8, je však ještě rozšířeno o prosté prahování.

1.2.3 EPIPOLÁRNÍ GEOMETRIE A TRIANGULACE

V předchozích odstavcích byly zmíněny možné cesty k extrakci významných bodů, jako například rohy. Jejich detekce je obzvláště důležitá pro následné zpracování v 3D rekonstrukci. Bude-li pro odhad bodu ve scéně X z bodu x ze snímku, byl by tento úkol, kvůli převodu 3D snímku na 2D nemožný (obr. 1.10 (a)).

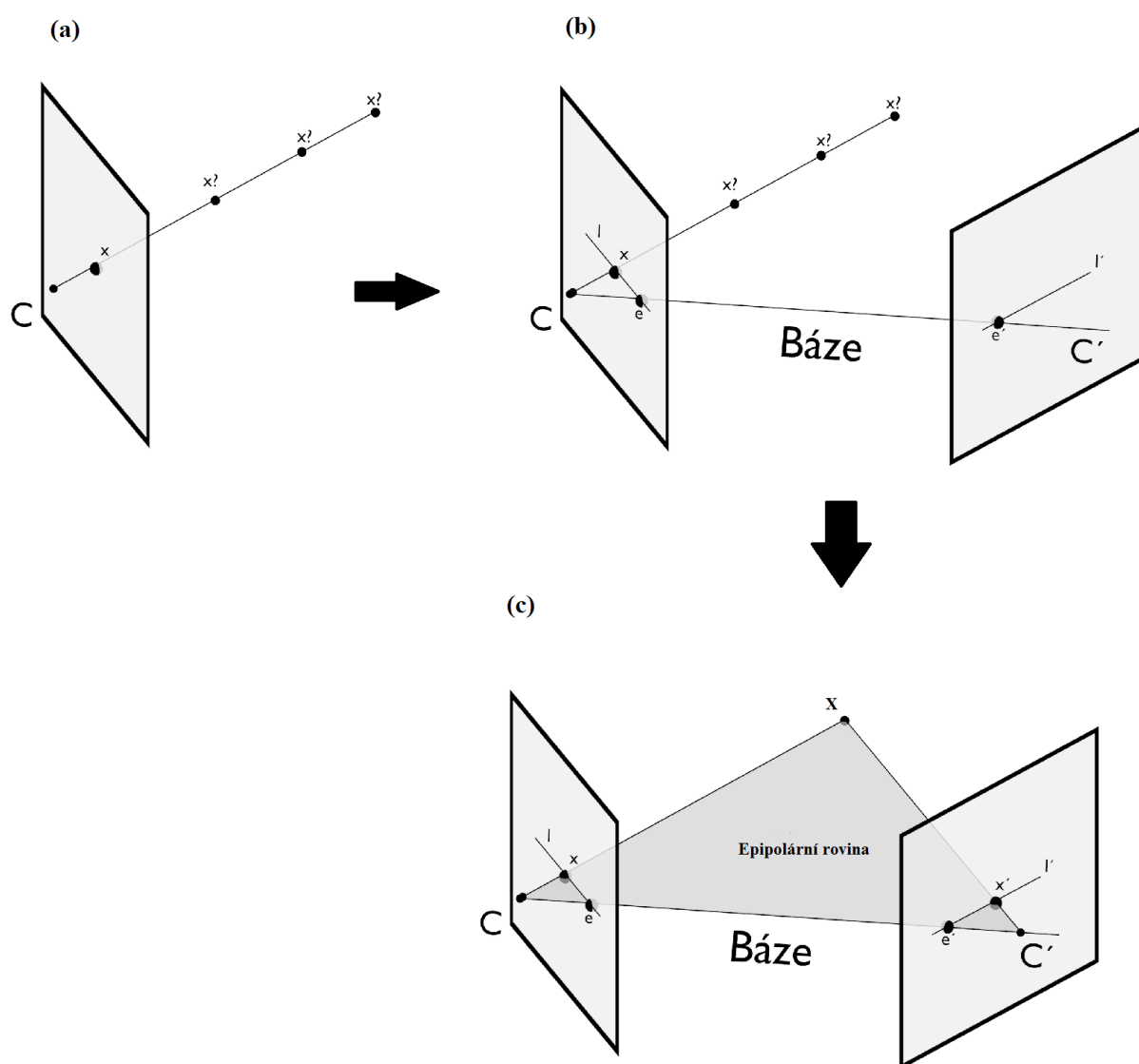
Bude-li však k dispozici další snímek, v případě další kamera, C' , bude tento úkol lépe realizovatelný (obr. 1.10 (b)). Problém se tak z 2D redukoval na 1D, respektive na linii. Pomocí dvou a více pohledů lze tak, přes triangulaci, dopočítat 3D pozici bodu X ve scéně [6].

Dříve, než bude problematika probrána dále, bude nutné si vymezit pár pojmů, které budou použity v následujícím výkladu: epipolární přímka (zkráceně epipoláry; l, l') vzniká průnikem epipolární a obrazové roviny, pomocí epipolární přímky je problém vyhledávání korespondujícího bodu z prvního snímku na druhý omezen na jednu přímku. Epipól (e, e'), je bod, který vznikl průnikem přímky, procházející mezi středy promítání (kamer) a obrazovou rovinou, v epipólu se protínají epipolární přímky. Dále zmíněný pojem je tzv. báze, jedná se o vzdálenost mezi kamerami (případně snímky). Středů promítání, C, C' a zkoumaný bod ve scéně, X , tvoří tzv. epipolární rovinu (π).

¹³ Průhlednost je zde šachovnicové pozadí, nicméně se lze setkat i s černým či bílým pozadím.

Jak již bylo zmíněno, pomocí dalšího pohledu (C') došlo k jistému vymezení. Nabízí se nicméně dvě otázky, jak nalézt korespondující bod v druhém snímku, respektive jeho správnou pozici? A jakou pozici má druhý snímek (případně kamera) vůči prvnímu. Jinými slovy, je potřeba promítnout bod x na epipolárnu (l') druhého snímku. Touto problematikou se zabývá tzv. epipolární geometrie.

Jedná se o koplanární geometrickou projekci mezi dvěma pohledy. Je nezávislá na struktuře scény, závisí pouze na vnitřních parametrech kamer a jejich vzájemném postavení [19]. Nalezení korespondence bodu je realizováno projekcí bodu x z prvního snímku na epipolární přímku snímku druhého buď pomocí esenciální anebo fundamentální matice.



Obr. 1.10: Problematika epipolární geometrie ilustrována na dvou pohledech (ve stereoskopické rekonstrukci na dvou kamerách). C a C' jsou středy projekce. (a) značí primární problém pro odhad hloubky (respektive vzdálenosti) bodu ve scéně, (b) vymezení problému na 1D, kde bod x je promítán na epipolární přímce. (c) projekce bodu x na epipolární přímku l' .

ESENCIÁLNÍ MATICE

Jak již bylo zmíněno, esenciální (a fundamentální) matice, E , umožňuje mapovat bod z prvního snímku x na epopoláru l' u snímku druhého, označený, jako x' . Je definována [5]:

$$(x'_n)^T E x_n = 0 \quad (1.24)$$

, x_n a x'_n jsou normalizované obrazové souřadnice vzájemných bodů u obou snímků [19]. Normalizovanými souřadnicemi jsou souřadnice scény orientovanými vůči kameře, body se tak neorientují vůči světovým souřadnicím, ale vůči kameře. Esenciální matice se využívá, je-li kamera kalibrována a obsahuje 5 parametrů, které obsahují informaci o translaci a rotaci vůči kameře (3 parametry jsou rotační a 2 translační).

FUNDAMENTÁLNÍ MATICE

Pro nalezení vzájemnosti bodu z prvního pohledu na druhý jsou využity pixelové souřadnice [19]. Matematická formulace je stejná, nicméně obrazové souřadnice nejsou normalizované [5]:

$$(x')^T F x = 0 \quad (1.25)$$

Vztah mezi fundamentální a esenciální je vyjádřen:

$$E = K'^T F K \quad (1.26)$$

, kde K a K' je matice vnitřních parametrů kamery, E je esenciální matice a F je matice fundamentální [19].

TRIANGULACE

Pomocí triangulace lze vypočítat pozici korespondujících bodů v prostoru, respektive bodu X . Prerekvizitou je epipolární geometrie, díky které lze zmíněná korespondence bodu určit. Jsou-li známy korespondující body x a x' a matice kamer P a P' , lze prostorový bod X vypočítat pomocí [19]:

$$x = PX, x' = P'X \quad (1.27)$$

1.3 REKONSTRUKČNÍ ALGORITMY

Je k dispozici mnoho 3D rekonstrukčních přístupů. Lze je rozdělit na dvě kategorie, na metody aktivní a pasivní, přičemž tyto kategorie lze ještě rozdělit na metody jedнопohledové a více pohledové.

Aktivní techniky 3D rekonstrukce jsou založené na interakci osvětlení s pozorovaným objektem. Může se jednat o laser anebo světlo, většinou strukturované do černých a bílých pásků. Aktivní techniky využívá řada 3D scannerů. Jejich hlavní výhodou je možné využití za špatných světelných podmínek (špatné osvětlení, podvodní průzkumy, zejména pak v zakalených vodách) [7], jinými slovy, nejsou příliš závislé na kvalitě snímku. Nevýhodou je

nicméně náročnost na vybavení, jelikož kromě fotoaparátu, respektive kamer, jsou potřeba další zařízení, jako například laser, LED anebo projektor. Další limitující skutečností je výstup objektu bez textury, jelikož, jak již bylo zmíněno, nejsou založené na kvalitě textury, nýbrž na interakci objektu s osvětlením.

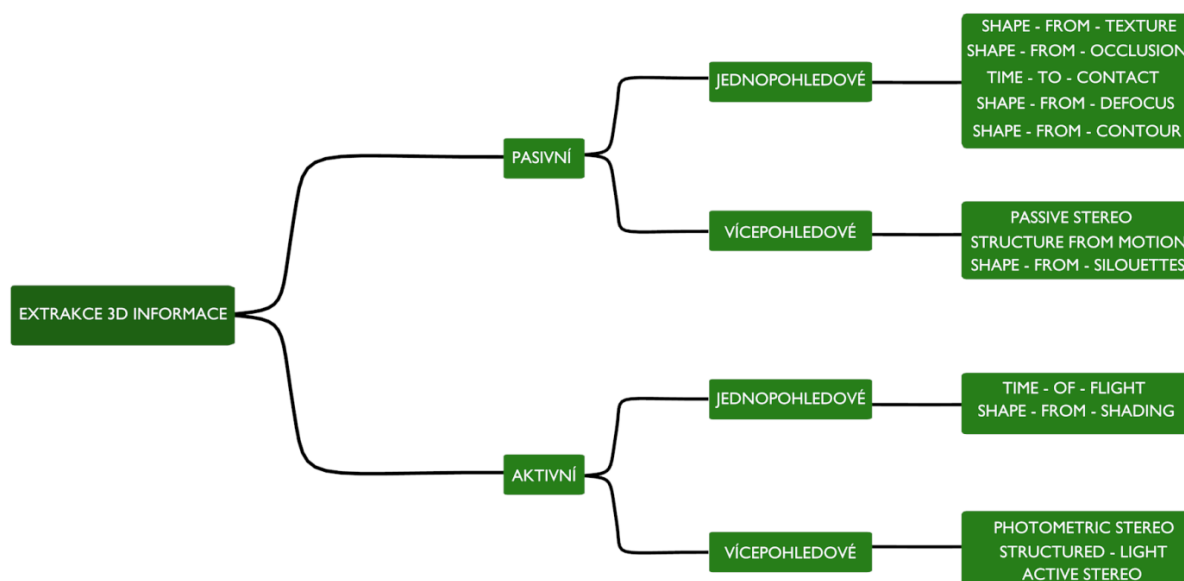
Pro druhou skupinu, pasivní metody, je charakteristická nenáročnost na vybavení, jelikož tyto techniky pracují pouze se získanými snímky, tudíž je vybavení tvořeno jedním anebo více fotoaparáty (případně kamerami). Jelikož tyto techniky pracují se snímky, jsou rovněž založené na kvalitě snímků, což zpravidla bývá faktor limitující kvalitu výstupu. Další výhodou, oproti metodám aktivním, je možnost 3D rekonstrukce objektu s vytexturovanými povrchy. Světlo zde slouží pouze jako osvětlovací prostředek.

Další rozdělení je založeno na počtu zachycených (příp. i osvětlených) stran objektu [6]. Jedn pohledové metody pracují pouze z jednoho bodu pohledu (dokonce i z jednoho snímku), metody více pohledové s vícero stranami objektu.

Taxonomie rekonstrukčních metod, tedy rozdělení do jednotlivých kategorií a výpis základních rekonstrukčních algoritmů, je znázorněna na Obr. 1.11.

Z obrázku 1.11 je patrné, že rekonstrukčních přístupů je mnoho, nicméně existuje jich mnohem více. Pro základní algoritmy platí jistá omezení v jejich využití, jelikož volba jednoho z těchto přístupů závisí na vlastnostech objektu, tedy, jeho geometrii, materiálu a na jiných charakteristikách. Tento problém řeší další rekonstrukční přístupy, tzv. hybridní techniky, které vznikají kombinací jednotlivých algoritmů z obr. 1.11. Cílem této hybridizace je dosažení co nejvyšší univerzálnosti, jelikož některé objekty jsou možné rekonstruovat ve 3D pouze pomocí některých algoritmů, zatímco v jiných případech se pro rekonstrukci hodí metody jiné. Například metoda „shape-from-shading“ je vhodná pro objekty, které jsou texturově chudé (jednobarevné, obtížné na detekci význačných bodů), zatímco pasivní „shape-from-texture“ metoda pracuje s pravidelnými texturami. Z výše uvedeného vyplývá, že při vhodném návrhu a kombinaci technik realizace pro multiskopickou 3D rekonstrukci je možné využít dokonce i jedn pohledové metody (komplikované), případně kombinací aktivních a pasivních metod. Lze tak s jistotou říci, že programátor má jistou „volnost“, co se týče využití rekonstrukčního přístupů, přičemž kombinací jednotlivých technik lze tak vytvářet nové algoritmy.

První část této kapitoly bude věnována procedurám (např. detekci rohů, triangulací...) nutným k realizaci 3D rekonstrukčních algoritmů, přičemž některé algoritmy budou popsány v poslední části kapitoly.



Obr. 1.11: Základní taxonomie metod pro extrakci 3D informace zkoumaného objektu, zde jsou uvedeny základní přístupy, přístupů nicméně existuje více, jako například tzv. hybridní, které vznikají kombinací těchto základních technik. Adaptováno z [6].

SHAPE-FROM-SHADING

Ačkoliv se jedná o jednopohledovou techniku, je možné ji využít i v přístupech multiskopických. Například kombinací s více pohledovou stereorekonstrukcí (Multi – view stereo) lze vytvořit algoritmus pro 3D rekonstrukci, kterého by samotné algoritmy nedocílily. Zatímco více pohledová stereorekonstrukce pracuje efektivně s vytexturovanou scénou, není schopen zachytit vysokofrekvenčními detaily, případně je citlivý na stínování způsobené nevhodným osvětlením rekonstruovaného objektu. Naopak jednopohledový „Shape-from-shading“ využívá stínování, nicméně jeho samotné použití pro více pohledovou 3D rekonstrukci je velmi obtížné (nikoliv nemožné). Jejich kombinací lze dosáhnout velmi kvalitních výsledků [16].

Metoda využívá vztahu, mezi úhlem dopadajícím a odraženém od povrchu objektu a normálou plochy, jedná se tak o metodu aktivní. Nejen, že je potřeba znát zdroje světla, ale také odrazivostní vlastnosti povrchu [6].

STEREOREKONSTRUKCE

Metoda využívá dvou kamer, které bývají synchronizovány [6]. Možné využití je v robotice, kdy tato technika nahrazuje pohled očí, robot je tak schopen získat informaci o vzdálenosti objektu. I přesto, že je metoda binokulární, lze tuto metodu aplikovat více pohledově („multi-view stereo“). Jedná se o hloubkovou 3D rekonstrukci, kdy výstupem je hloubková mapa na základě které je rekonstruovaná scéna a odhadována vzdálenost objektů. Je založena na

detekci významných bodů u dané dvojice snímků. Následuje nalezení korespondence pomocí epipolární geometrie. Dalším krokem je tzv. rektifikace, což je proces, kdy jsou snímky zarovnány paralelně k bázi. Výstupem stereorekonstrukce je tzv. disparitní mapa, jak moc se snímky po rektifikaci posunuly buď doprava anebo doleva k druhému snímku. Disparitní mapa je nepřímo úměrná s hloubkovou mapou. Čím je tedy větší disparita, tím nižší je hloubka [2][5][6].

STRUCTURE-FROM-MOTION

Jedná se o techniku, která na základě dvou a více snímků na vstupu, odhaduje 3D strukturu na výstupu. Výhodou této metody je možné využití bez potřeby kalibrace. Odhad parametrů kamery si provádí algoritmus sám na základě obrazových příznaků. Tato auto-kalibrace je nicméně možná pouze tehdy, pokud se jedná o jednoduché lineární transformace, jako je translace a rotace. Tudíž je realizace možná i s kalibrací kamery [21]. Realizace je podobná, jako při stereorekonstrukci, tedy jsou založené na detekci významných bodů, nalezení vzájemných bodů, výpočet jejich souřadnice v prostoru. Hlavní rozdíl oproti stereorekonstrukci je fakt, že zatímco u stereorekonstrukce je využito dvou kamer, je u metody „structure from motion“ (dále jen SfM) využito pouze jedné kamery, přičemž je tento přístup založený na snímání (a rekonstrukci) objektu za pohybu. Dalším rozdílem je využití SfM bez použití kalibrace [5][6]. Jak již bylo zmíněno, realizace je mírně podobná stereorekonstrukci. Je provedena detekce významných bodů, například pomocí Harris-Stephensonova algoritmu. Dále se využívá deskriptorů pro uchování informace příznaků v invariantní informaci. Následuje nalezení korespondujících bodů a určení jejich pozice v prostoru pomocí triangulace. Nakonec je výstup interpretován, jako mračno rozmístěných korespondujících bodů v prostoru (v případě stereorekonstrukce je to hloubková mapa) [2][5][6].

STRUKTURA ZE SILUET („SHAPE-FROM-SILHOUETTES“)

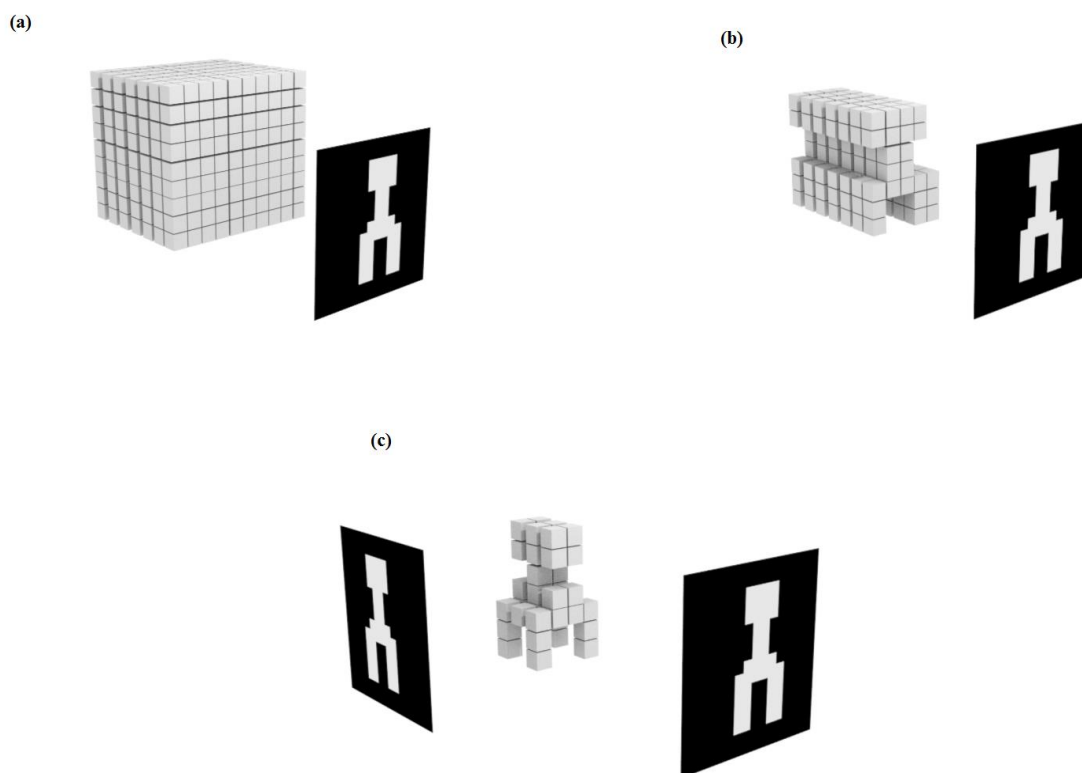
Jedná se o pasivní vícehledovou rekonstrukční metodu, která, jak již práce napovídá, pracuje se siluetami. Siluety zde reprezentují binární obraz, kde většinou 1 reprezentuje objekt a 0 jeho pozadí. Vytvoření siluet bylo vysvětleno v sekci o segmentaci obrazu, zde bude pouze popsán samotný algoritmus.

Siluety lze chápat, jako binární snímek, kde logická hodnota 1 je vysegmentovaný objekt zájmu a 0 vše okolo.

Využitím vícero snímků siluet téhož objektu je intersekcí projekcí možné rekonstruovat objekt. Metodu lze uplatnit pro rekonstrukci objektů s hladkým povrchem (čož by jiné, texturově založené, přístupy nedokázaly, případně velmi obtížně), rovněž je tento přístup schopný „real-time“ rekonstrukce. Je k dispozici několik přístupů.

Jedním z možných přístupů je tzv. **voxel carving**, doslovně přeloženo, jako řezba voxelu. Pro 3D rekonstrukci je využitý objekt tvořený voxely. Siluety zde představují masku, ta interaguje s objektem. Problematiku si lze představit následovně: Je-li silueta postavena před objektem, pak všechny voxely, které jsou za siluetou (bílé pixely), jsou zachovány. Naopak

voxely, které se nachází za nevysegmentovanou částí (mimo siluetu - za černými pixely), jsou odstraněny. Objekt byl tak vytvářen z jedné strany. Je-li použito více siluet, zobrazující různé strany objektu, lze objekt pomocí siluet upravovat až do konečné podoby. Ilustrativní ukázka metody siluet je na obrázku 1.12.



Obr. 1.12: „Shape-from-silhouette“ pomocí voxel carving, (a) voxelový objekt a silueta na vstupu, (b) objekt upravený interakcí se siluetou, (c) postupným přidáváním dalších siluet (respektive pohledů) lze objekt dále upravovat až do požadované podoby.

využitím kalibrace lze dosáhnout přesnějších výsledků. Lze totiž vzít v potaz perspektivní projekci, čímž se částečně změní podoba rekonstruovaného objektu a zamezí se nepřesnostem [26][28].

Dalším možným rekonstrukčním přístupem je tzv. **visual hull**. Jedná se o přístup založený na intersekcí vizuálních kuželů, tvořených projekcí siluet do 3D prostoru. Jedná se o kalibrační přístup. Toho se dá docílit buď prostřednictvím výše zmíněného (voxel carving s využitím kalibrace kamery) anebo využitím epipolární geometrie, kde se nehledá korespondence mezi body, ale mezi tzv. zjevnými konturami (z angl. apparent contours). Jedná se o příznaky zakřivených objektů (v případě binárních snímků se tak jedná o okraje siluet) [23][27].

Přívětivějším realizačním způsobem výše zmíněných přístupů je rekonstrukce napodobením tomografických medicínských systémů (CT, MRI...) přes **prostou zpětnou projekci**. Jedná se o metodu, která spočívá v rozprostření všech binárních snímků (siluet) přes rekonstrukční matici pod úhlem, pod kterým byly jednotlivé snímky získány [8].

2 NÁVRH

V rámci bakalářské práce byly vybrány techniky „structure-from-motion“ (dále jen SfM) a „shape-from-silhouettes“ (dále jen SfS), pomocí zpětné projekce.

Obě metody budou rekonstruovat skutečné objekty z 25-ti snímků (foceno 360° s krokem 15°). Pro porovnání, byly, u obou modalit, pro 3D rekonstrukci použity stejné objekty. V případě metody SfS se ještě před tímto krokem provede rekonstrukce virtuálních objektů. Taktéž budou snímány 360° , nicméně s krokem 1° . Bude tak pro rekonstrukci použito 361 snímků. Důvody volby nižšího kroku pro snímání virtuálních objektů budou podrobeny diskuzi. V případě SfM bylo pro snadnější určení vnitřních a vnějších parametrů kamery využito kalibrace.

Práce byla zhotovena v programovém prostředí *MATLAB 2018b*, využitím příkazů z knihovny funkcí *Image Processing Toolbox™* a *Computer Vision System Toolbox™*, který poskytuje algoritmy pro návrh a simulaci počítačového vidění, čili je možné provést detekci a extrakci příznaků, nalezení shod, kalibraci kamery, 3D rekonstrukci a zpracování 3D mračna bodů [11].

2.1 METODA „STRUCTURE FROM MOTION“ (SfM)

Primárním krokem bylo vytvoření akvizičního systému. Ten se skládal z digitálního fotoaparátu Nikon D5200 s teleskopickým objektivem Nikon AF-S DX NIKKOR 55-200 mm f/1.4-5.6G ED, stativem a otočným stolem.

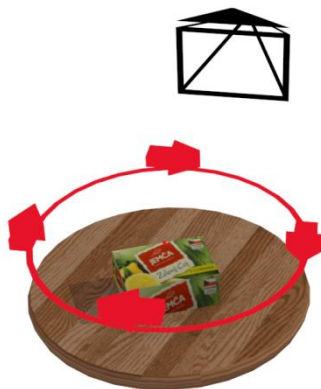
V ideálním případě by na každý pohled objektu měla být jedna kamera, čímž by vznikla složitá soustava kamer. Tyto kamery by měly objekt nasnímat ve stejnou chvíli, aby pohledy objektu byly nasnímány za stejných podmínek (zejména, je-li objekt pohyblivý). Skutečnost je však taková, že realizace takového systému je finančně a realizačně náročná. Je proto třeba zvolit kompromis, který umožní taktéž rekonstruovat objekt z více stran, ale bude vyžadovat méně kamer. Rizikem je však možné zhoršení přesnosti výstupu.

Na základě technických možností autora (disponuje jen jedním fotoaparátem) a výše zmíněné skutečnosti se nabízejí dva akviziční přístupy.

V prvním případě by objekt stál na místě a pohybovala by se kamera. S využitím kalibrace kamery by to ale znamenalo, že by byla třeba neustálá výměna snímaného objektu s kalibrační šachovnicí, čímž může dojít k nepřesnostem. Tento problém lze zredukovat využitím druhého přístupu. Nepohybovala by se kamera, nýbrž objekt. Pohyb bude však rotační a to kolem vlastní osy, čímž bude zachována stejná vzdálenost objektu od fotoaparátu a bude možné provést 25 snímků z různých stran objektu, přičemž postačí jen jedna kalibrace kamery. Pro názornost byl na obr. 2.1 znázorněn navržený akviziční systém. V této práci bude využitý druhý přístup, kde rotaci objektu kolem vlastní osy bylo využito pomocí rotačního stolku.

Po získání parametrů kamery a snímků objektu bylo možné pokračovat v realizaci 3D rekonstrukce. Nejdříve bylo nutné napravit fotografie postihnuté distorzí čoček. Následovala detekce příznaků, extrakce a vyhledání jejich vzájemností mezi snímky. Dále se využila

epipolární geometrie a následoval výpočet pohybu kamery, čímž bylo umožněné pokračovat ve výpočtu 3D lokace shodných bodů pomocí triangulace. Nakonec se ze vstupních snímků získala barva pro výstupní mračna bodů, reprezentující realizovaný objekt (dále point cloud). Tato mračna bodů se nakonec spojila.



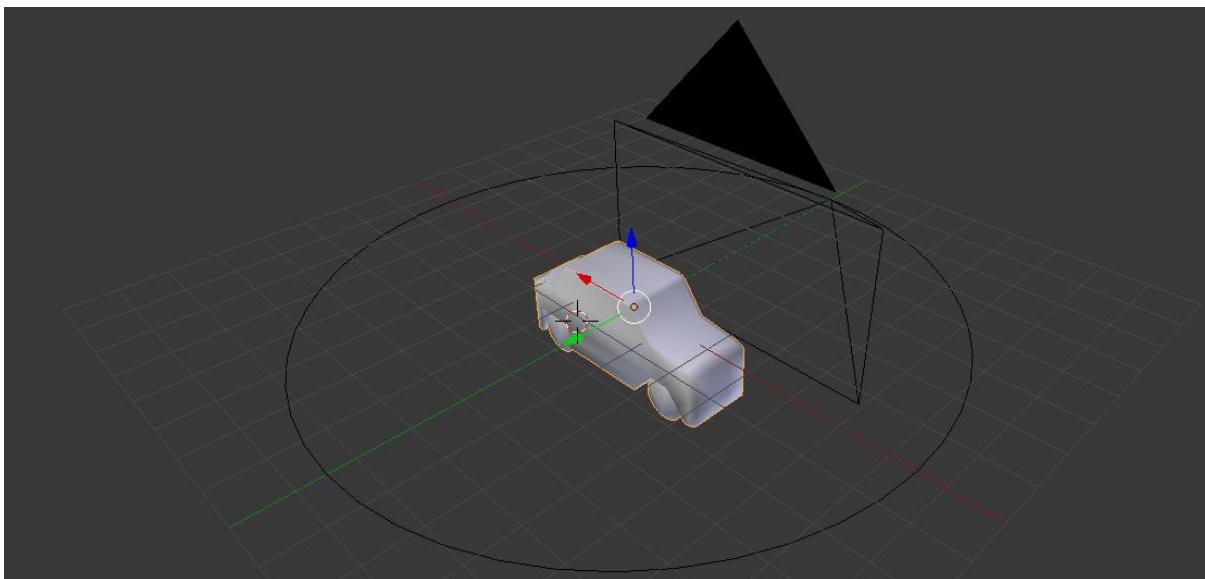
Obr. 2.1: Návrh akvizičního systému

2.2 METODA „SHAPE FROM SILHOUETTES“ (SfS)

V případě druhé techniky, SfS, se provedl pokus, jak na snímcích reálného, tak na snímcích virtuálních objektů. Pokus na snímcích virtuálního objektů se volil, jako první. Ty se vymodelovaly v programu Blender 2.78. Jedná se o open-source program, určený pro 3D počítačovou grafiku. Má mnoho funkcí, nicméně je zejména využíván pro 3D modelování, simulace, tvorbu a úpravu (nejen) animovaných videí, tvorbu animací, případně může také plnit funkci, jako herní engine.

Jelikož se algoritmus realizoval přes tomografickou zpětnou projekci, nebyla třeba kalibrace. Nicméně i přesto se musel zvolit vhodný postup, který by napodoboval tomografické zobrazovací systémy (CT, MRI apod.). Bylo třeba brát na vědomí, že snímky daného objektu musely být zhotovené ze stejné vzdálenosti. Rovněž bylo třeba se vyhnout při renderování snímků perspektivní projekci, jelikož u ní platí nepřímá úměra mezi vzdáleností objektu od kamery a velikostí pixelů v obraze, což nebylo žádoucí. Byla tedy pro zhotovení snímků potřeba ortografická projekce.

V případě programu Blender postačovalo zvolit možnost ortografická projekce. Dále, aby byl objekt nasnímán ze stejné vzdálenosti, bylo třeba si vytvořit pomocnou křivku, konkrétně kruh, jehož střed byl rovněž středem objektu. Po této křivce se pohybovala kamera, jejíž ostřicí bod byl rekonstruovaný objekt. Zhotovila se animace o délce 361 snímků, kdy jeden snímek reprezentoval posun kamery o jeden úhel (tedy od 0 do 360 stupňů). Následně se provedlo renderování animace, bylo nicméně třeba před provedením vykreslování nastavit, aby místo .AVI formátu provedlo renderování snímek po snímku ve formátu .PNG, čímž se tak získalo 361 snímků. Návrh akvizičního systému v programu Blender je na Obr.2.2.



Obr. 2.2: Návrh akvizčního systému v programu Blender

Pro získání snímků skutečného objektu bylo třeba zrealizovat podobný akvizční systém. Ten byl, stejně jako u akvizčního systému pro SfM tvořen fotoaparátem Nikon D5200, stativem a otočným stolem, byla ale použita jiná orientace kamery (bez natočení, stejná orientace, jako na Obr. 2.2).

Jelikož zde byla perspektivní projekce nežádoucí, bylo potřeba snímky zhotovit tak, aby alespoň částečně napodobovaly výstup z ortografické projekce. Toho se docílilo volbou větší ohniskové vzdálenosti. Využil se proto opět teleskopický objektiv Nikon AF-S DX NIKKOR 55-200 *mm* f/1.4-5.6G ED.

Oproti vykreslování virtuálního objektu se netočila kamera, ale otočný stolek. Postup byl tedy téměř stejný, jako u metody SfM. Bylo opět potřeba zachovat stejnou vzdálenost mezi objektem a kamerou. Pokud by se pohyboval fotoaparát, tak nejen, že by tento postup byl zdlouhavý, ale rovněž by vzniklo riziko, že by vzdálenost kamery od objektu byla různá, což by mělo velmi výrazný vliv na konečný výstup. Otáčením objektu by se problém částečně redukoval, nicméně i zde by bylo stále riziko, že by se otočením objektu mírně posunul a došlo by tak opět ke změně vzdálenosti. V případě, otáčení kruhového stolku, se problém redukoval na minimum. Po přesném zvolení středu, na který se položil objekt, bylo zaručeno, že nejen stolek, ale také objekt se bude točit kolem své vlastní osy, čímž bylo zaručeno zachování stejné vzdálenosti mezi objektem a fotoaparátem. Ilustrativní znázornění návrhu akvizčního systému pro získání snímků skutečného objektu je na Obr. 2.1.

Získané snímky se následně nahrály a převedly na šedotónové, následně na binární obrazy. Dále se postupovalo tak, že se vytvořily rekonstrukční a tři pomocné matice (3D). Do pomocné matice se rozprostřel řez po řezu snímek o daném úhlu. Následně se jednotlivé řezy matice (2D matice) otočily o úhel, pod kterým byl určitý snímek zhotoven. Tato modifikovaná matice se následně přičetla k rekonstrukční matici. Postup byl iterativně aplikován i u ostatních snímků. Nakonec se prostřednictvím prahování získal rekonstruovaný objekt. Konkrétní realizace algoritmu bude zmíněna v další kapitole.

3 REALIZACE

Následující kapitola se věnuje konkrétnímu popisu realizace jednotlivých pokusů v prostředí *MATLAB 2018b*, které byly zmíněny v předchozí kapitole. Jejich výsledky budou podrobeny diskuzi v následující kapitole.

3.1 METODA „STRUCTURE FROM MOTION“

První zvolenou metodou byla metoda structure from motion (SfM), založená na kalibraci kamery. Následující podkapitola rozebere podrobnější postup jednotlivých kroků realizace tohoto algoritmu. První část budou tvořit nutné kroky před samotným programovým řešením rekonstrukční metody, tedy převzorkování snímků a kalibraci. Následně bude popsán postup realizace samostatného algoritmu, tedy detekce a extrakce významných bodů, aplikace epipolární geometrie a triangulace, obarvení mračen bodů a nakonec jejich spojení do jednoho, finálního, mračna bodů.

3.1.1 PŘEVZORKOVÁNÍ SNÍMKŮ

Expozice snímků byla manuální, bez blesku se shodnými vlastnostmi, tedy: nastaveném rozlišení 6000 x 4000 pixelů, 300dpi, bitové hloubce 24bit, ohniskové vzdálenosti 55 mm, ISO – 1600, rychlostí expozice 1/2500 sec. a při maximální cloně 4.

Vysoké rozlišení snímků by bylo na zpracování příliš náročné, bylo proto nutné snímky převzorkovat pomocí funkce `imresize` na rozlišení 600 x 400 pixelů.

3.1.2 KALIBRACE KAMERY

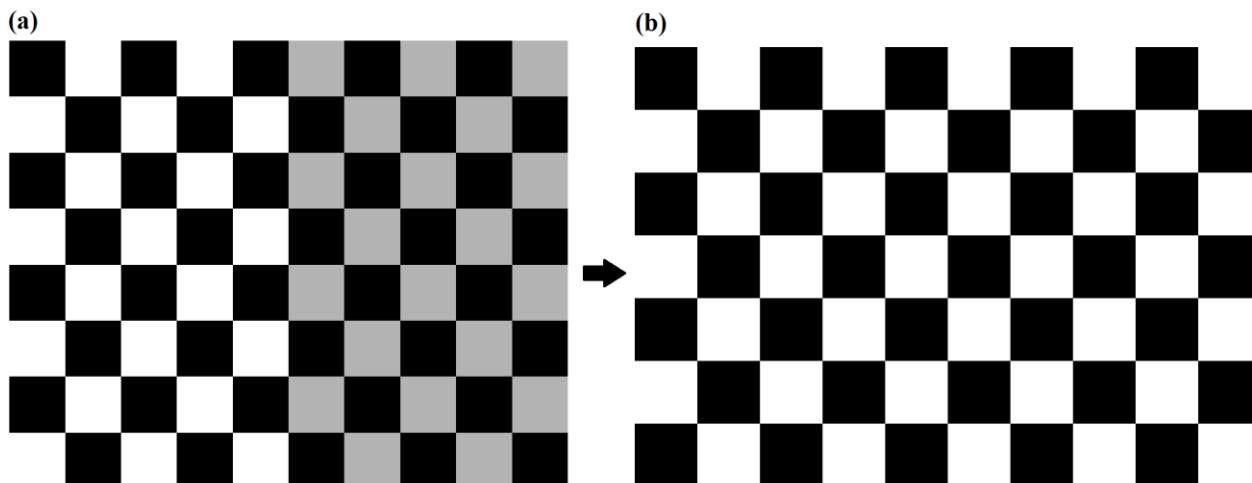
Pro kalibraci kamery byla použita aplikace *Camera Calibrator App*, vytvořený Jean-Yves Bouguetem, který využívá Zhangovu kalibrační metodu [6][14][15].

VYTVOŘENÍ KALIBRAČNÍHO OBJEKTU

Ačkoliv bylo v kapitole o kalibraci kamery řečeno, že lze využít různé kalibrační objekty, má *Camera Calibrator App* jistá kritéria pro provedení správné kalibrace.

Pro správné výsledky vyžaduje černobílou šachovnici, jejíž dílky sice musí být čtvercové, nicméně jako celek musí mít tvar obdélníku, kde na jedné straně musí být sudý a na druhé lichý počet černých čtverečků. Dále musí být splněna podmínka, aby na jedné straně byly v rozích pouze černé čtverce a na straně protilehlé naopak bílé. Delší strana šachovnice je zde chápána, jako osa x . Tato kritéria nicméně umožňují aplikaci odhadovat orientaci šachovnice v prostoru (bez ohledu na orientaci os kalibračního objektu).

Z důvodu přesnosti byla vytvořena vlastní šachovnice, v prostředí MATLAB ji lze vygenerovat pomocí příkazu `checkerboard(n, a, b)`, kde n byl rozměr dílku v pixelech, a počet černých dílků na výšku, b dílků na šířku. Vygenerovaná šachovnice nejen, že neměla požadované rozložení, ale také pravá polovina šachovnice nebyla černobílá, nýbrž černošedá (obr. 3.1 (a)). Taková šachovnice by na kalibraci kamery byla nevhodná. Problém se vyřešil prostým oříznutím a prahováním obrazu (obr. 3.1 (b)).



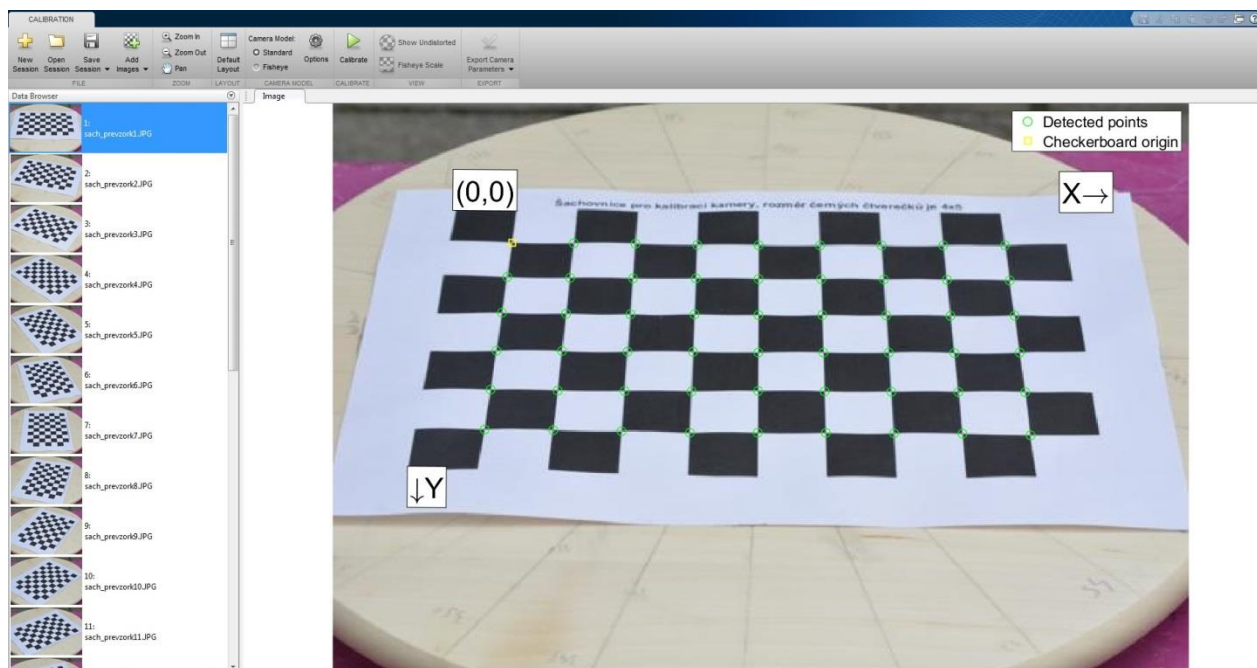
Obr. 3.1: Vygenerovaná šachovnice pro kalibraci kamery. (a) první podoba šachovnice, která je nicméně pro kalibraci nevhodná, (b) konečná podoba šachovnice, splňující požadavky pro Single Camera Calibrator App, rozměr jednoho dílku je 24 mm.

POSTUP KALIBRACE KAMERY

Vytvořením šachovnice bylo možné pokračovat dále. Pro lepší výsledky dle MathWorks® doporučeno zhotovit 10-20 snímků kalibrační šachovnice z různých pohledů, přičemž je vhodné je snímat pod určitým úhlem (doporučeno méně, než 45°), nejméně však aplikace vyžaduje tři.

Dále bylo nutné, aby vzdálenost šachovnice od kamery byla stejná, jako vzdálenost kamery od rekonstruovaného objektu. Rovněž je (v případě fotoaparátu) použita manuální expozice a deaktivován autofocus.

V rámci bakalářské práce byla použita kalibrační šachovnice s rozměry čtverců 24 mm, která byla zaznamenána z různých pohledů. Bylo zhotoveno celkem 25 snímků. Ty byly následně nahrány do *Camera Calibrator App*. Po nahrání snímků je tato aplikace schopná detekovat rohy na šachovnici. Ukázka prostředí s nahranými kalibračními snímky je na obrázku 3.2.



Obr. 3.2: Ukázka prostředí Camera Calibrator App s nahranými snímky kalibrační šachovnice (před spuštěním kalibrace).

Dalším krokem bylo spuštění kalibrace. Ta se spustí kliknutím v horním panelu na zelenou šipku s textem calibrate. Před spuštěním je nejdříve nutné zadat velikost jednoho dílku. Byla zadána hodnota 24 mm. Kalibrace se následně provede samostatně.

Po dokončení kalibrace se v prostředí zobrazí výsledky. Ty tvoří již detekované body a jejich reprojekce (zpětná projekce) pomocí získaných parametrů kamery. Rovněž byl k dispozici náhled odhadu vnějších parametrů kamery (Obr. 4.2 (b)) a reprojekčních chyb (Obr. 4.2 (a)). Chyba vzniklá zpětnou projekcí je dána vzdáleností mezi detekovaným bodem a bodem vzniklým jeho reprojekcí. Platí zde tedy přímá úměra, tedy, čím větší vzdálenost těchto bodů, tím menší chyba nastala.

Nakonec se parametry kamery exportovaly do „worksapce“ v MATLAB a uložily pro budoucí práci.

3.1.3 REALIZACE PROGRAMU

Prvním (a zásadním krokem) algoritmu bylo načtení parametrů kamery a vstupních snímků. Parametry kamery byly nahrány pomocí příkazu `load`. Snímky byly nahrány pomocí příkazu `imread`, přičemž získané (přečtené) informace z těchto grafických souborů se uložily prostřednictvím `for` cyklu do buňkového pole `cell`. Buňkové pole tak obsažovalo 25 snímků. Pro zobrazení snímku lze použít příkaz `imshow`, nicméně lze zobrazit i dvojici snímků pomocí `imshowpair`, kde je použita metoda 'montage' anebo lze použít příkaz `montage` pro zobrazení vícero obrázků.

Tyto snímky se opět prostřednictvím `for` cyklu pro snadnější práci převedly na snímky šedotónové a uložily do buňkového pole `gray`.

Pro vyšší přesnost rekonstrukce bylo vhodné přes `undistortImage` napravit případné chyby vzniklé distorzí čoček. Vstupy příkazu jsou snímek a parametry kamery, které kromě vnitřních a vnějších parametrů obsahují i distorzní koeficienty, díky kterým lze napravit případné chyby (obrázek 4.4). Jak si lze však z obrázku 4.4 (b) povšimnout, že díky použití vyšší ohniskové vzdálenosti ($f = 55$ mm), nebyly vstupní snímky příliš postiženy distorzí.

Následoval samotný rekonstrukční algoritmus realizovaný prostřednictvím for cyklu, kde se hledání korespondencí, epipolární geometrie a triangulace prováděla se sousedním snímkem. Postupovalo se tak po dvojicích. Jelikož budou následující kroky pro všechny dvojice snímků stejné, budou tyto kroky popsány jen pro jednu z těchto dvojic.

DETEKCE A EXTRAKCE VÝNAMNÝCH BODŮ

Dále kroky směřovaly k detekci významných bodů, díky kterým bylo možné nalezení korespondencí. Byl využitý příkaz `detectHarrisFeatures` pro nalezení rohů, na základě Harris-Stephensonova algoritmu, který byl zmíněn v první kapitole.

Vstupem příkazu byl analyzovaný obraz a možné argumenty, které funkci více specifikují. Jako například `'MinQuality'`, který vyjadřuje minimální akceptovanou kvalitu pro detekci rohů. Čím nižší je hodnota, tím vyšší výtěžek z detekce (na úkor vyšší nepřesnosti). V rámci pokusu bylo, z důvodu zisku více detekovaných bodů, navolena povolená minimální kvalita na defaultní hodnotě (0,01), čímž se sice získalo méně bodů, ale detekce byla tak přesnější.

Jak již bylo v kapitole o detekci rohů zmíněno, pro nalezení korespondujících bodů je nutné zajistit, aby byl nalezen tentýž bod. Je proto nutné, aby tento bod byl, z hlediska informace, neměnný například na úhlu pohledu anebo na různém osvětlení. Toho lze docílit pomocí `extractFeatures`, který zajistí extrakci deskriptoru bodu. Vstupem je snímek a nalezené rohy. Dále je možné, jako vnitřní argument, si vybrat descriptor dle potřeby. Zde byl vybrán deskriptor BRISK¹⁴.

EPIPOLÁRNÍ GEOMETRIE A TRIANGULACE

Pro nalezení korespondujících bodů byla použita fundamentální matice přes `estimateFundamentalMatrix`. Jako výstupy vyžadovala hrubě nalezené shody, ty se zjistily přes `matchFeatures`, kde výstupem byla matice obsahující index korespondujících párů. Jako metoda pro výpočet fundamentální matice byl použitý RANSAC.

Pro uskutečnění triangulace se využil příkaz `triangulate`. Na vstupu nicméně vyžadoval výpočet projekční matice prvního a druhého snímku.

Přes `relativeCameraPose` se vypočítala pozice druhého snímku vůči prvnímu. Vstupem byly výstupní hodnoty z `estimateFundamentalMatrix` a parametry kamery. Výstupem byl translační vektor (rozměr 1x3) a rotační matice (rozměr 3x3).

¹⁴ Binary Robust Invariant Scalable Keypoints, zkráceně BRISK, je rychlejší alternativou deskriptorů SURF či SIFT [25]

Projekční matice u první kamery byla brána jako referenční, tudíž pro výpočet přes `cameraMatrix` byl translační vektor roven `[0, 0, 0]` a rotační matice byla rovna nulové matici o rozměru `3x3`. Nakonec se z relativní pozice kamery vypočítala matice kamery druhého pohledu využitím `cameraPoseToExtrinsics`, čímž se určil vnější parametr druhého pohledu, a tak se získala prostorová orientace pozic kamery u obou pohledů.

Pomocí triangulace (přes příkaz `triangulate`) se vypočítaly 3D souřadnice korespondujících bodů. Ty bylo následně nutné převést do interpretující formy. Toho se docílilo skrze mračna bodů funkcí `pointCloud`, kde vstupem byly nejen již zmíněné 3D souřadnice korespondujících bodů, ale také barva (zmíněno dále). Nakonec se jednotlivá mračna bodů sjednotila a zobrazila pomocí `pcshow` a uložila do buňkového pole `temp`. Ten se využil, jako dočasná proměnná uchovávající všechna mračna bodů, která se následně sjednotila.

OBARVENÍ MRAČNA BODŮ

Jelikož jsou jednotlivé body point cloudu tvořeny korespondujícími body, bylo možné pro získání barvy zvolit jeden z nich. Nicméně byl zvolen první z dvojice, jelikož i při výpočtu pozice druhého snímku byl první snímek volen, jako referenční. Nejdříve bylo potřeba vytvořit proměnnou obsahující všechny body s informací o jejich barvě. Bodů může být až $v \times 3$ = šířka \times výška. Pomocí tří for cyklů byla vytvořena proměnná `varcolor` o rozměrech $v \times 3$. Matice je tvořena řádky, přičemž sloupce reprezentují složky RGB barevného modelu. Jinými slovy, první sloupec obsahuje barevnou složku R (red), druhý sloupec složku G (green) a třetí barevnou složku B (blue). Následně bylo potřeba obarvit pouze body, které byly detekovány, proto se vytvořila proměnná `Index`, které obsahovala pouze indexy detekovaných bodů (pomocí `sub2ind`, kde vstupy tvořil počet řádku a x/y souřadnice korespondujících bodů). Na konec se vytvořila proměnná `color`, která indexaci proměnné `varcolor(Index)`.

SJEDNOCENÍ MRAČEN BODŮ

Bylo řečeno, že jednotlivá mračna bodů se spojila do jednoho celku. Toho se docílilo prostřednictvím aplikace transformací a jejich registrací.

Geometrické transformace lze rozdělit na lineární a nelineární. Mezi lineární transformace patří například otočení (rotate), posunutí (translace) a změna měřítka (scaling). Nelineární transformace řeší složitější změny objektu, například deformaci [1]. Na zmíněná mračna bodů budou použity lineární transformace, konkrétně rotační matice. Ty jsou pro jednotlivé osy matematicky popsány, jako [1]:

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$R_y(\alpha) = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \alpha & 0 & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.1)$$

$$R_z(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 & 0 \\ \sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Tyto transformační matice bylo možné vytvořit pomocí příkazu `makehgtform`, kde vstupem byl argument, udávající rotaci kolem určité osy a úhel v radiánech. V případě rotace kolem osy-z byl argumentem textový řetězec `'zrotate'`.

Následně se jednotlivé rotační matice mezi sebou vynásobily a uložily do proměnné `R` reprezentující rotační transformace okolo všech os. Tato proměnná se použila, jako vstup do funkce určenou pro transformaci mračna bodů `pctransform`, kde vstupem je vedle již zmíněné transformační matice mračno bodů. Tento výstup se následně účastnil registrace mračna bodů prostřednictvím funkce `pcregisterndt`. Získaný výstup se použil pro transformaci druhého pointcloudu, kde opět pomocí funkce `pctransform` se výstup z registrace mračna bodů použil jako transformační matice. Druhé mračno bodů, které se transformovalo, se následně spojilo s prvním pomocí příkazu `pcmerge`. Výstup se uložil do proměnné `ptCloudFin`. Stejný postup pro registraci následující point cloudy prostřednictvím `for` cyklu, kde roli prvního mračna bodů zastupovala již proměnná `ptCloudFin`. Nakonec se tato proměnná zobrazila pomocí příkazu `pcshow`, čímž se vizualizovala konečná podoba výsledného mračna bodů, kde, pro přesnost, se následně využilo příkazů `xlim`, `ylim` a `zlim`, které omezily rozsahy os.

3.2 METODA „SHAPE FROM SILHOUETTES“

Druhým zvoleným rekonstrukčním přístupem, který byl podroben pokusu, byla metoda `shape from silhouettes` (SfS), založenou na tomografické prosté zpětné projekci. Následující podkapitola rozebere podrobnější postup jednotlivých kroků realizace tohoto algoritmu.

Jak již bylo zmíněno v předchozí kapitole, pokus byl realizován jak na skutečných objektech, tak i na objektech virtuálních. Jelikož programové řešení bylo stejné, bude popsán postup obecně, bez potřeby rozdělení realizace algoritmu na objekty reálné a virtuální.

3.2.1 PŘEVZORKOVÁNÍ SNÍMKŮ

Z důvodu vysokého rozlišení a většího počtu snímků bylo nutné všechny snímky ve `for` cyklu převzorkovat pomocí funkce `imresize` u virtuálních objektů na 1/20 a reálných objektů na 1/15 svých původních rozměrů.

3.2.2 PŘEVOD BAREVNÝCH SNÍMKŮ NA ŠEDOTÓNOVÉ

Pro snadnější práci bylo nutné všechny snímky převést na šedotónové obrazy. Z důvodu většího počtu snímků se opět využil `for` cyklus, ve kterém se ale využil jiný příkaz, `rgb2gray`.

3.2.3 VYTVOŘENÍ SILUET

V obou případech, tedy u virtuálního a skutečného souboru, bylo nutné vytvořit siluety. Ty se vytvořily pomocí nejjednoduššího způsobu segmentace, prostého prahování. Nutno zmínit, že z důvodu nerovnoměrného umístění u snímků skutečného objektu, bylo potřeba využít i úprav v programu Adobe Photoshop CS6, ve kterém se získané snímky zarovnaly a odstranilo se pozadí.

Prahování se provedlo tak, že se nastavil práh, zde $T = 0.1$. Následně se přes cyklus `for` provedlo prahování tak, že veškeré jasové hodnoty pixelů aktuálního (*i*-tého) obrazu, které jsou vyšší, než práh, budou nabývat hodnoty 1. Zbytek bude roven hodnotě 0. Výstupem byl tedy binární obraz. Pro následující práci matlab nicméně vyžadoval, aby získané výstupy byly v datovém typu `double`. Proto se prahování ošetřilo tak, že se prahování vložilo do funkce `double`, která výstup po prahování převedla na tento datový formát.

3.2.4 REKONSTRUKCE STRUKTURY ZE SILUET

Získáním siluet bylo možné pokračovat dále v rekonstrukci. Vytvořili se tři pomocné matice (které byly prázdné) a nulová rekonstrukční matice, `RekMatice`, o rozměrech $[y, x, z]$. Hodnoty x, y reprezentují rozměry obrázků. Ty se získaly pomocí příkazu `size`, kde vstupem byl libovolný snímek účastníci se rekonstrukce (všechny snímky měly stejné rozměry). Poslední rozměr, z , byl roven maximální hodnotě z dvojice x, y , který se získal příkazem `max`.

Následoval `for` cyklus, jehož počet iterací (a jejich krok) byl založen na počtu úhlů (a kroku), pod kterými byly jednotlivé snímky pořízeny. Ten tvořil samostatnou rekonstrukci, v němž se nacházely další cykly `for`. První, v něm obsažený cyklus, o počtu iterací $1:z$, sloužil pro uložení *i*-tého snímku do pomocné matice, `Mat1`, do všech jeho řezů po ose z . Tato pomocná proměnná posloužila ve druhém cyklu ($iterace = 1 : y$). Ten iterativně vykonával následující postup: *i*-tý řez se uložil do druhé pomocné proměnné `Mat2`. Tato proměnná reprezentovala řez, respektive 2D obrázek. Proto bylo potřeba využít příkaz `squeeze`, aby se zachoval 2D rozměr. Kdyby se nevyužil, vznikl by vektor, nikoliv matice, což by bylo nežádoucí. Proměnná `Mat2` se následně otočila o určitý úhel (záleželo, který *k*-tý obrázek – úhel, byl v iteraci) pomocí `imrotate`, ve kterém se přidal argument `'crop'`, který výstupní snímek ořízl tak, aby měl stejné rozměry, jako snímek vstupní (`Mat2`). Následně se výstupní snímek (otočený a seříznutý) vložil do pomocné matice `Mat3`, do všech jeho řezů. Tato matice se následně přičetla do rekonstrukční matice `RekMatice`.

Proces se opakoval pro všechny snímky. Nakonec se objemová data z rekonstrukční matice extrahovala pomocí `isosurface`, čímž bylo možné 3D objekt zobrazit.

Konečná rekonstrukční matice byla tvořena sumou “rozkopírovaných“ siluet, kde společná oblast nabývala nejvyšší hodnoty. Rekonstrukce nicméně nebyla ještě hotová, bylo třeba z rekonstrukční matice odstranit nespolečné oblasti. Finální podoba rekonstrukce (Obr. 4.20, Obr. 4.22, Obr. 4.24, Obr. 4.26 a Obr. 4.28) se tedy provedla prahováním rekonstrukční matice, kde prahem byla jeho maximální hodnota (v případě snímků od 0° - 360° , s krokem 1° by se jednalo o hodnotu 361). Pro regulaci prahování byla tato hodnota násobena proměnnou `TreshLevel`, která reprezentovala procentuální úroveň této prahovací hodnoty. Hodnota 0 reprezentovala 0%, hodnota 1 naopak představovala 100% maximální hodnoty. Jinými slovy, v případě snímků od 0° - 360° , s krokem 1° a proměnné `TreshLevel` rovna jedné by se opět jednalo o hodnotu 361.

4 VÝSLEDKY

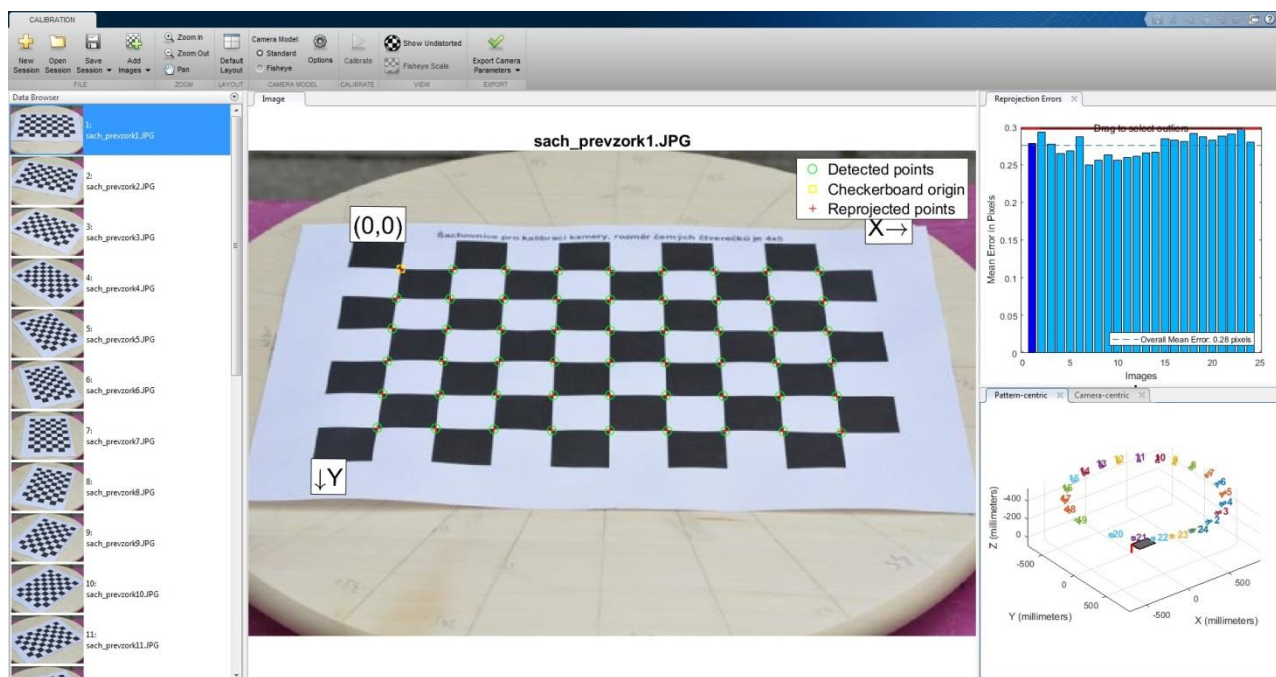
V této kapitole budou nejdříve zobrazeny výsledky, které budou v další kapitole podrobeny diskuzi. V případě metody SfM bude nejdříve zobrazen výstup kalibrace a následně zobrazené výstupy rekonstrukce u jednotlivých objektů.

4.1 METODA „STRUCTURE FROM MOTION“

V rámci této podkapitoly budou znázorněny výsledky rekonstrukce jednotlivých objektů využitím příkazu „structure from motion“. Podkapitola je rozdělena na 4 části. První část vyobrazuje výsledky kalibrace kamery, část druhá obsahuje výsledky 3D rekonstrukce krabičky od čaje, třetí rekonstrukci modelu bunkru a nakonec část poslední, čtvrtá, obsahuje získané výsledky rekonstrukce figurky vojáčka.

4.1.1 KALIBRACE KAMERY

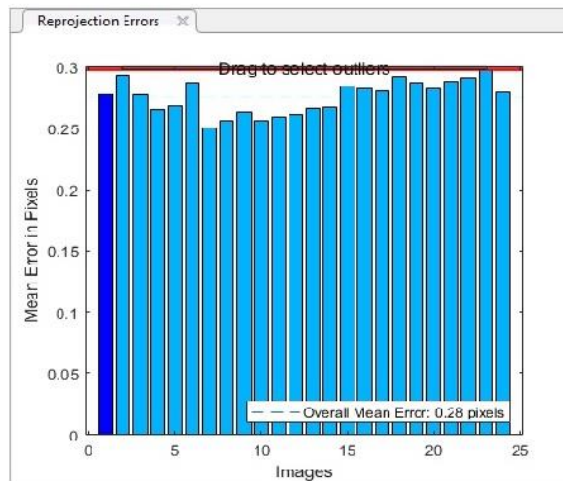
Spuštěním příkazu „kalibrate“ bylo dosaženo následujících výsledků:



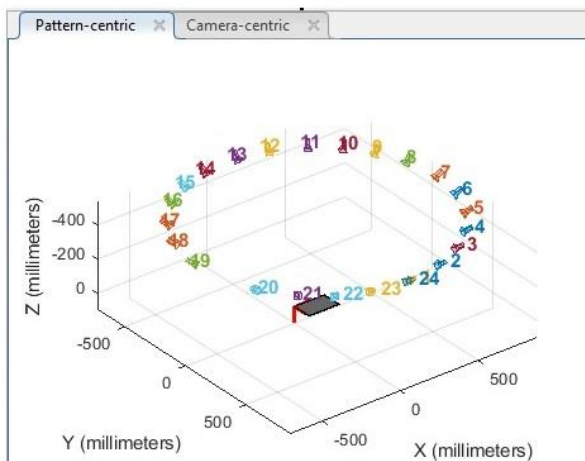
Obr. 4.1: Ukázka detekce rohu (zelená kolečka) a jejich reprojekce (červené křížky).

Z Obr. 4.1. je patrné, že snímků účastnících se kalibrace je 24, nikoliv 25. Poslední snímek (360°) byl identický se snímkem prvním (0°), byl proto z tohoto procesu odstraněn.

(a)



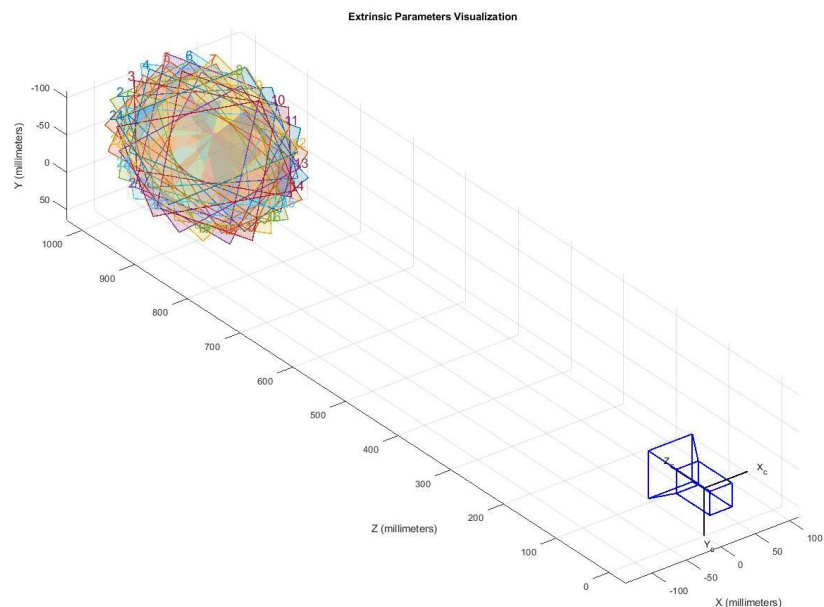
(b)



Obr. 4.2: (a) Bližší pohled na výsledky odhadu reprojekčních chyb, (b) odhad rozmístění kamery v prostoru

Na Obr. 4.2 (a) si lze povšimnout, že u 2. a 23. snímku nastala největší reprojekční chyba (v pravo nahoře, okénko „reprojection error“).

Nakonec se parametry kamery exportovaly do „worksapce“ v MATLAB a uložily pro budoucí práci.



Obr. 4.3: Odhadnuté vnější parametry kamery vizualizované v prostředí MATLAB

4.1.2 VZOREK Č. 1 - KRABÍČKA OD ČAJE

Z důvodu velkého počtu vstupů a výstupů, nebudou vyobrazeny všechny výsledky jednotlivých kroků (tedy například všechny detekce rohů, všechny páry mračen bodů apod.). Bude tak, ilustrativně, uvedena pouze vybraná část.

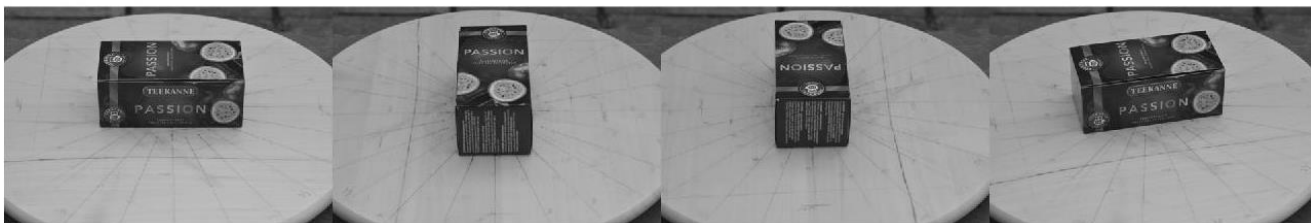
Na základě postupu, zmíněného v předchozí kapitole, bylo dosaženo následujících výsledků:

Prvním krokem bylo načtení vstupních snímků. Na Obr. 4.3 je názorná ukázka 4 vstupních snímků, reprezentující nasnímaný objekt z různých stran (pomocí otáčení stolku).



Obr. 4.3: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0°, (b) 90°, (c) 270° a (d) o 345°

Následujícím krokem byl převod RGB snímků na snímky šedotónové využitím příkazu `rgb2gray` (Obr. 4.4)



Obr. 4.4: Převedení barevných snímků na snímky šedotónové

Pomocí parametrů kamery bylo možné odstranit chyby u snímků, způsobené distorzí. Z důvodu použití vyšší ohniskové vzdálenosti, byly tyto chyby minimální (Obr. 4.5).



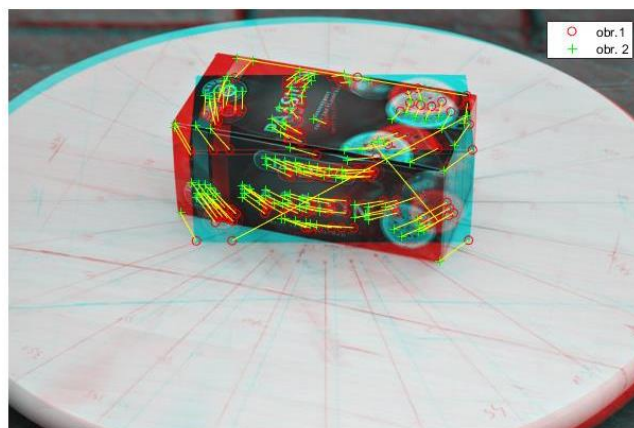
Obr. 4.5: Korekce snímků pomocí příkazu `undistortImage`

Následnou detekcí příznaků byly detekovány rohy. Na Obr. 4.6 je ukázka detekce rohů u prvního snímku



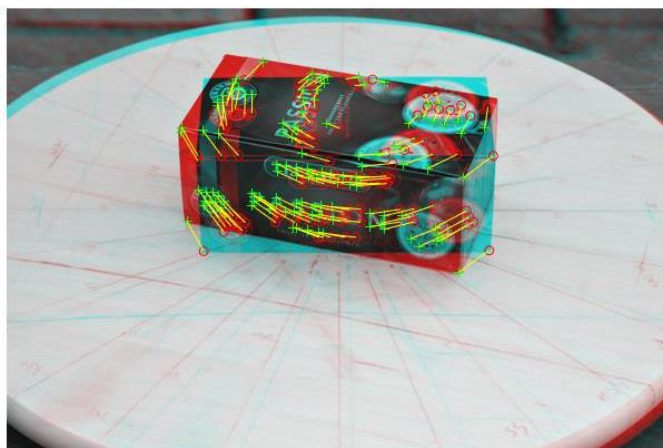
Obr. 4.6: Detekované rohy pomocí Harris-Stephensonova algoritmu u prvního snímku.

Detekcí příznaků obrazu bylo následně možné provést hrubé nalezení korespondujících bodů mezi jednotlivými dvojicemi snímků (Obr. 4.7).



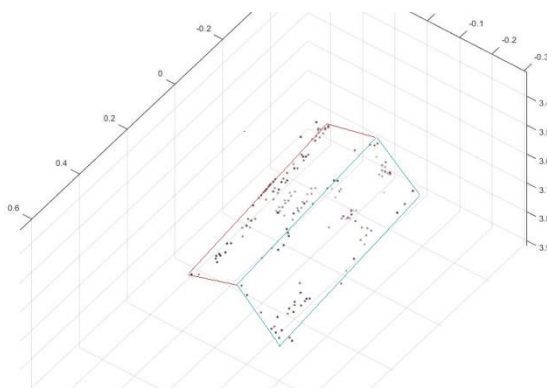
Obr. 4.7: Hrubé nalezení korespondujících bodů mezi první dvojicí snímků

Využitím epipolární geometrie se následně našly skutečné vzájemnosti (Obr. 4.8.), čímž se odstranily nesprávné vzájemnosti dané předchozím krokem.

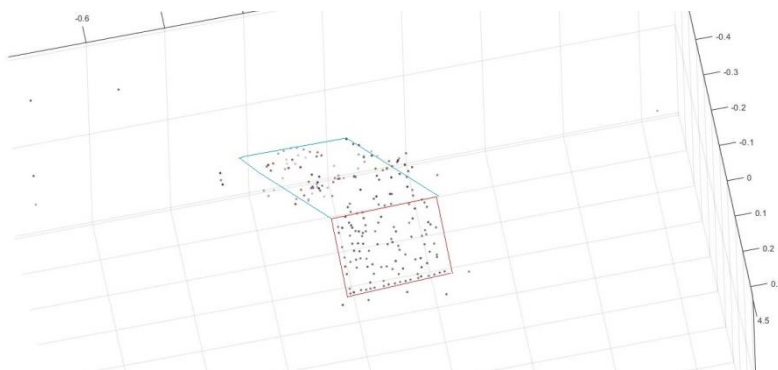


Obr. 4.8: Vymezení korespondujících bodů u první dvojice snímků

Následnými úpravami se rekonstruovala jednotlivá mračna bodů jednotlivých párů (Obr. 4.9. a Obr. 4.10).

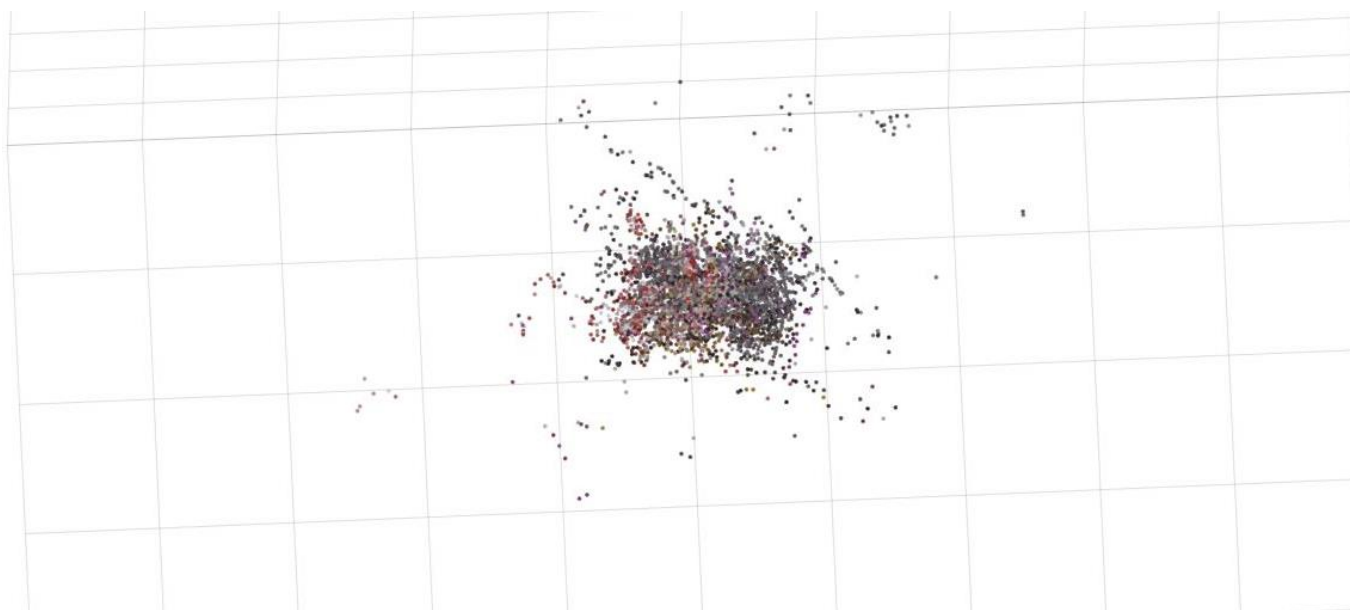


Obr. 4.9: 3D rekonstrukce první dvojice (natočení objektu o 0° a 15°)



Obr. 4.10: 3D rekonstrukce šesté dvojice (natočení objektu o 75° a 90°)

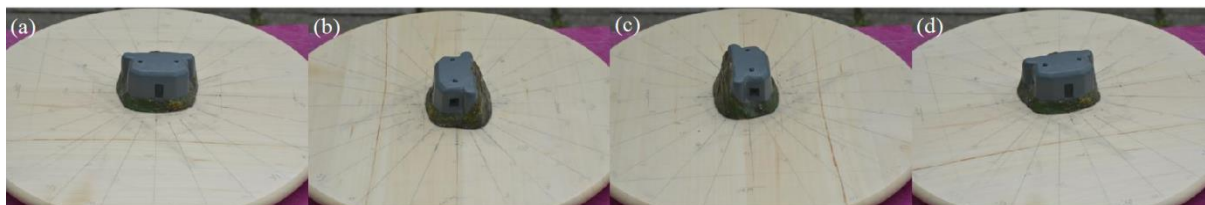
Nakonec využitím registrace mračen bodů a geometrických transformací bylo rekonstruováno konečné mračno bodů



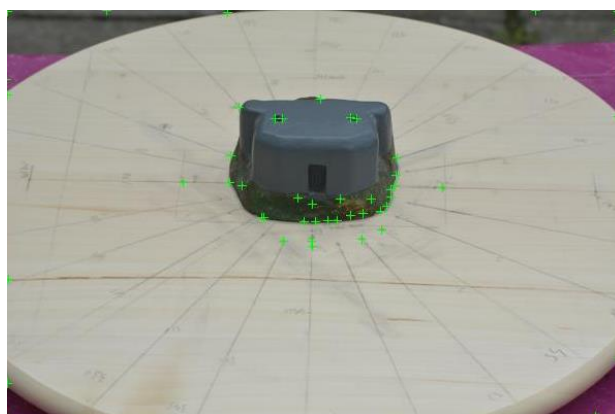
Obr. 4.11: Konečná podoba výsledného point cloudu krabičky čaje

4.1.3 VZOREK Č. 2 - SÁDROVÝ MODEL BUNKRU

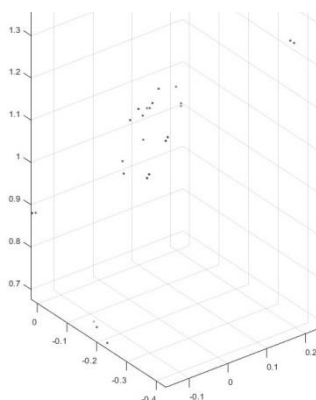
Jelikož se jednalo o texturově chudý objekt, nebylo možné detekovat dostatečné množství obrazových příznaků (Obr 4.13) ze vstupních snímků (Obr. 4.12), čímž se rovněž ovlivnilo nalézání korespondencí a celková rekonstrukce. Rekonstrukce tohoto objektu tak nebyla zdárná. Na Obr. 4.14 je ukázka rekonstrukce objektu u první dvojice snímků (pohledů).



Obr. 4.12: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°



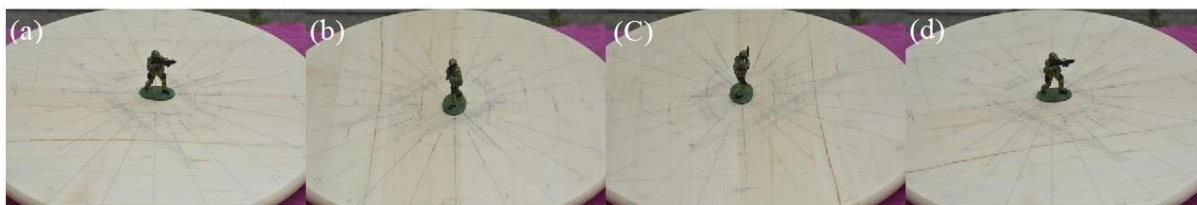
Obr. 4.13: Ukázka detekovaných rohů u prvního snímku



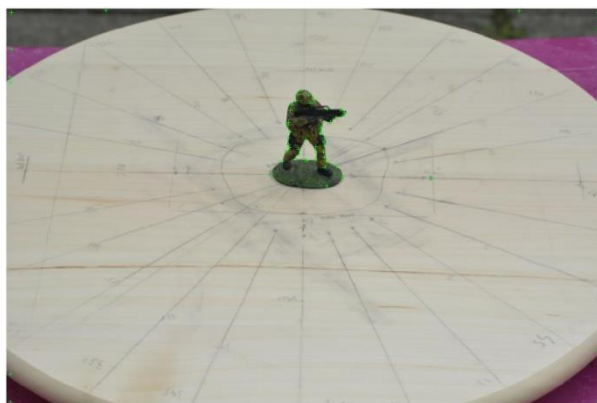
Obr. 4.14 : Neúspěšná rekonstrukce objektu dána vlivem nedostatečného množství nalezených korespondujících bodů. Zde se jedná o rekonstrukci první dvojice (natočení objektu o 0° a 15°).

4.1.4 VZOREK Č. 3 - MODEL FIGURKY VOJÁČKA

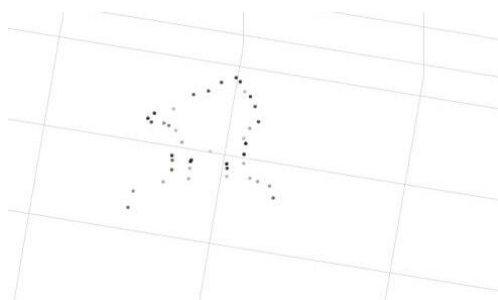
Stejně, jako u sádrového modelu bunkru, byl i tento objekt neúspěšně rekonstruován (Obr. 4.17). Malou velikostí bylo obtížné ze vstupních snímků (Obr. 4.15) detekovat dostatečné množství bodů (Obr. 4.16) a tak se program potýkal s podobnými problémy, jako u předchozího objektu. Rekonstrukce tohoto objektu taktéž nebyla úspěšná.



Obr. 4.15: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°



Obr. 4.16: Ukázka detekovaných rohů u prvního snímku



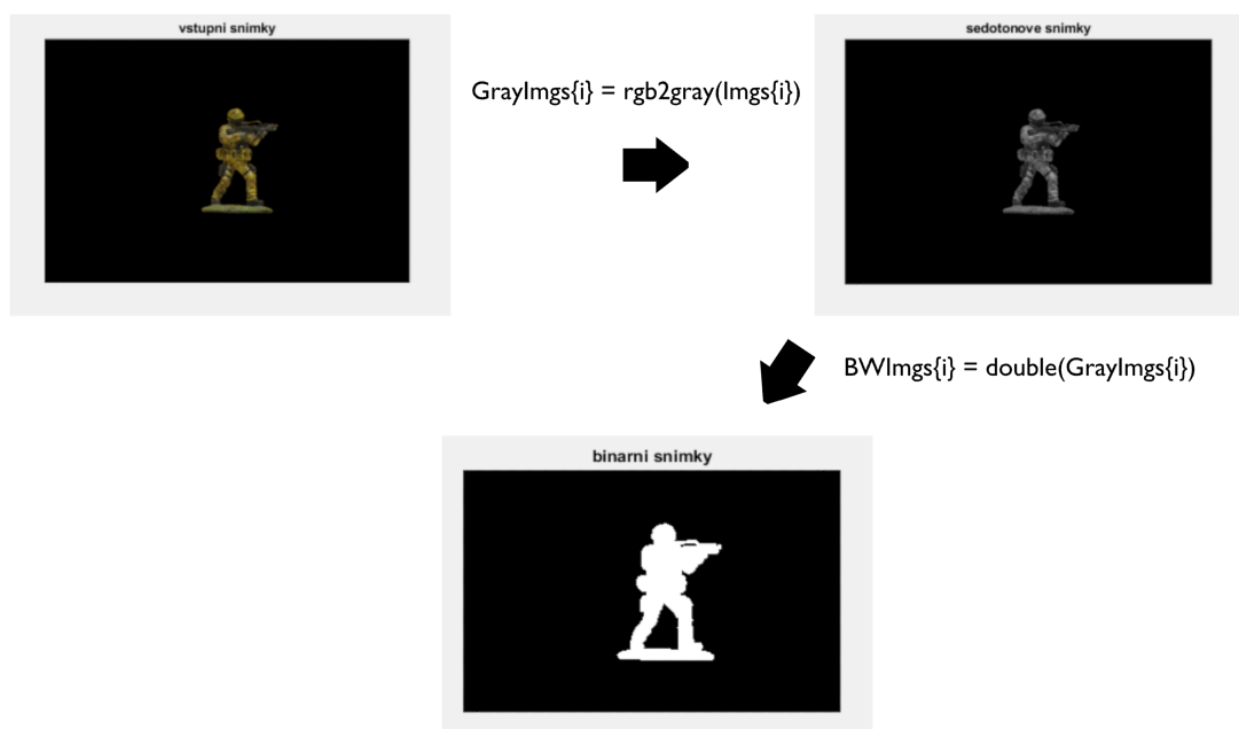
Obr. 4.17 : Neúspěšná rekonstrukce objektu dána vlivem nedostatečného množství nalezených korespondujících bodů. Zde se jedná o rekonstrukci druhé dvojice (natočení objektu o 15° a 30°).

4.2 METODA „SHAPE FROM SILHOUETTES“

V rámci této podkapitoly budou znázorněny výsledky rekonstrukce jednotlivých objektů, které se získaly aplikováním postupů zmíněných v předchozí kapitole. Podobně, jako u metody SfM bude tato podkapitola rozdělena na více částí. První část bude tvořit ukázka vytvoření siluet. Zbýlé části se budou již týkat výsledků 3D rekonstrukce jednotlivých objektů (virtuální, skutečné). Jelikož bylo během tohoto pokusu využito rozsáhlé množství snímků, budou jednotlivé výsledky reprezentovány pouze vybranou částí.

4.2.1 VYTVOŘENÉ SILUETY

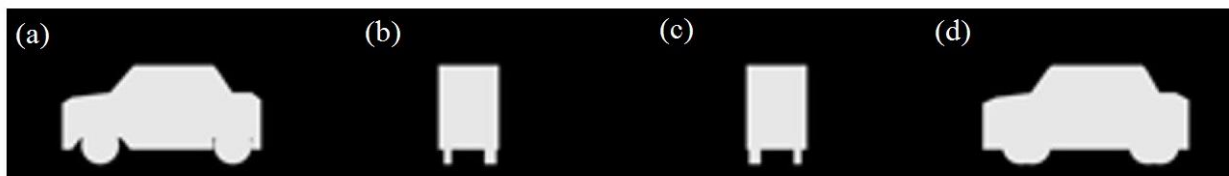
Převedením RGB snímku na šedotónový a následným použitím prostého prahování bylo dosaženo vytvoření siluety. Na Obr. 4.18 je ilustrativní ukázka vytvoření siluety na prvním snímku figurky vojáčka.



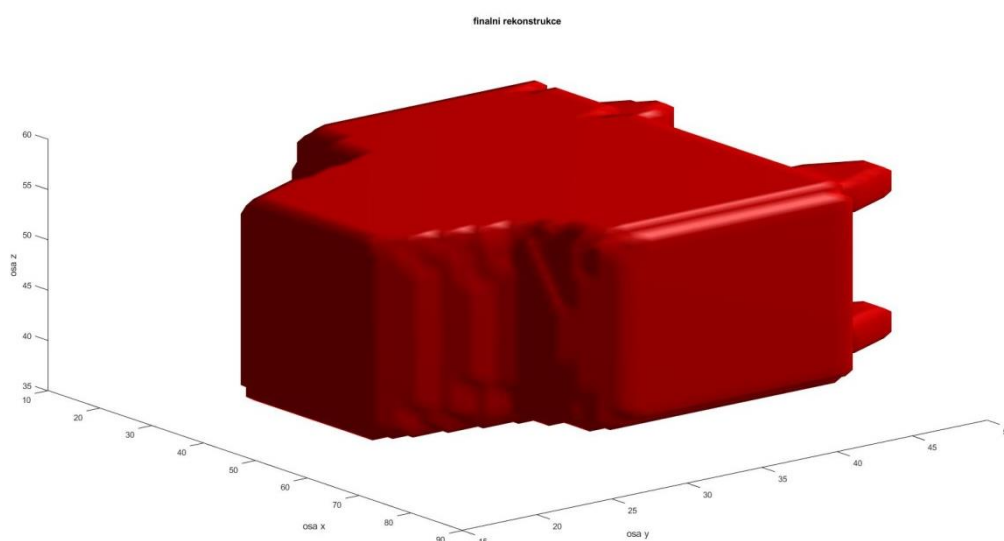
Obr. 4.18: Ilustrativní ukázka vytvořené siluety převedením RGB snímku na šedotónový a následným uplatněním prahování.

4.2.2 VZOREK Č.1 - VIRTUÁLNÍ MODEL AUTA

Za vstupních snímků (Obr. 4.19) objektu a aplikováním kroků zmíněných v předchozí kapitole bylo dosaženo výstupu, který je zaznamenán na Obr. 4.20.



Obr. 4.19: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0°, (b) 90°, (c) 270° a (d) o 345°



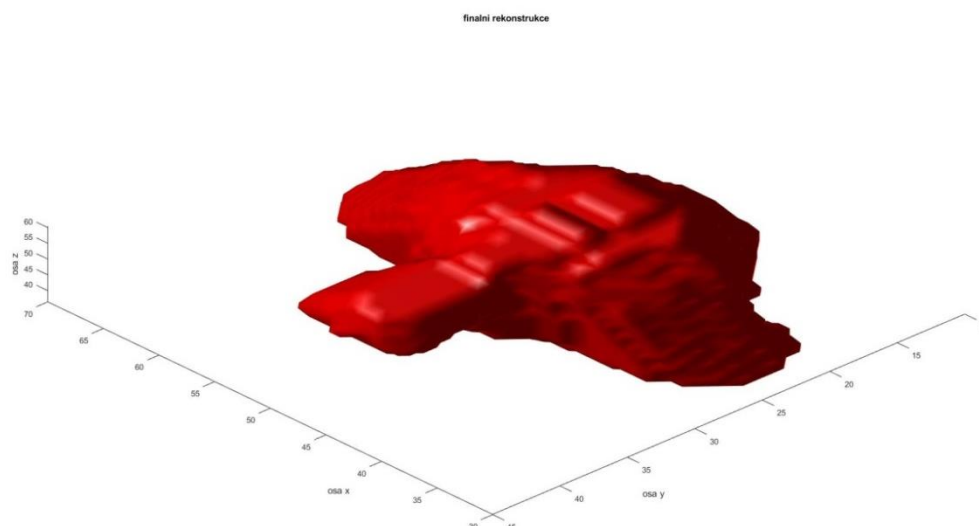
Obr. 4.20: Konečná rekonstrukce virtuálního modelu auta

4.2.3 VZOREK Č. 2 - VIRTUÁLNÍ MODEL OPIČKY

Za vstupních snímků (Obr. 4.21) objektu a aplikováním kroků zmíněných v předchozí kapitole bylo dosaženo výstupu, který je zaznamenán na Obr. 4.22.



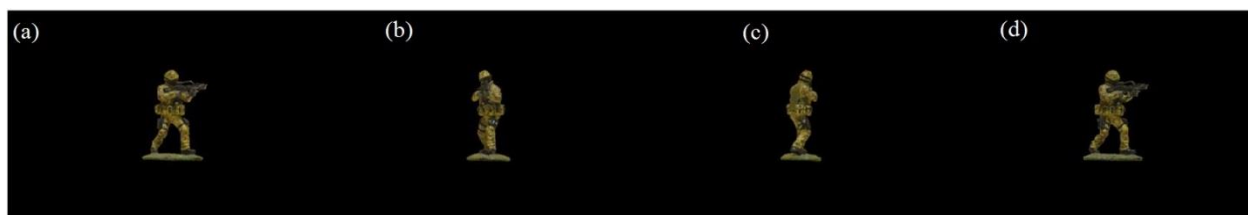
Obr. 4.21: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0°, (b) 90°, (c) 270° a (d) o 345°



Obr. 4.22: Konečná rekonstrukce virtuálního modelu opičky

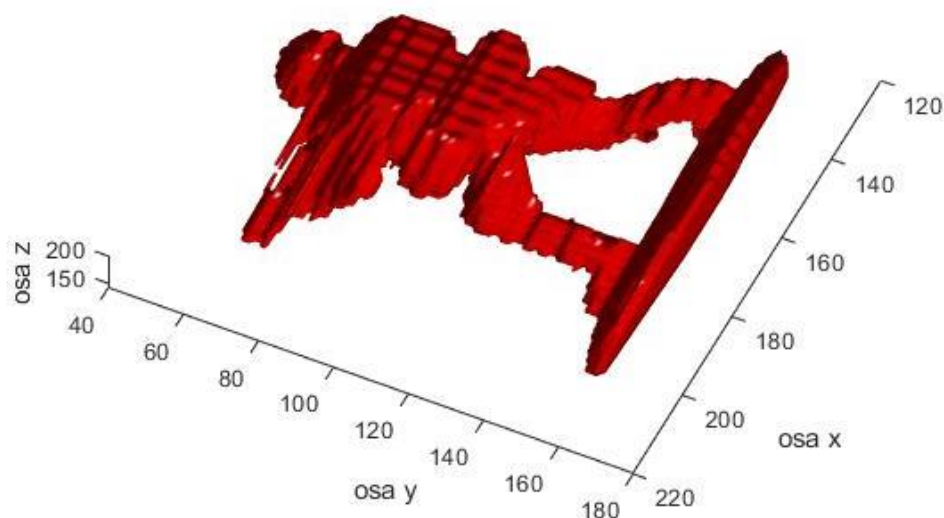
4.2.4 VZOREK Č. 3 - MODEL FIGURKY VOJÁČKA

Za vstupních snímků (Obr. 4.23) objektu a aplikováním kroků zmíněných v předchozí kapitole bylo dosaženo výstupu, který je zaznamenán na Obr. 4.24



Obr. 4.23: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

finalni rekonstrukce



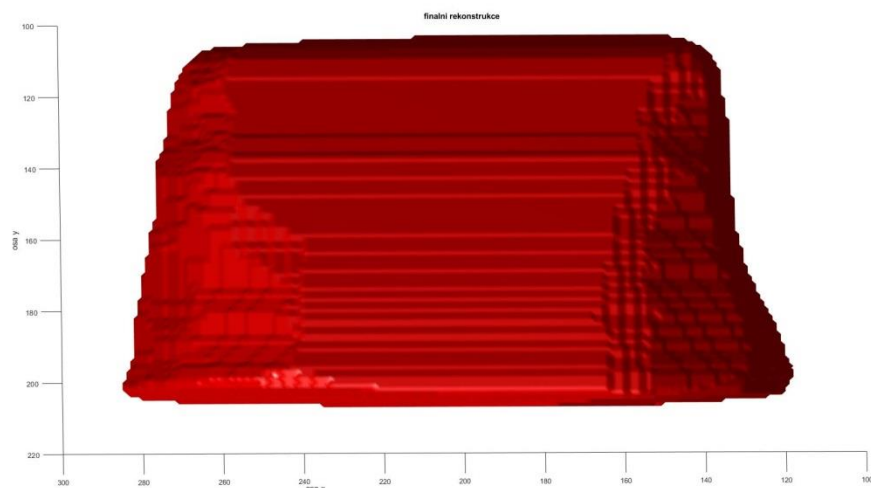
Obr. 4.24: Konečná rekonstrukce figurky vojáčka

4.2.5 VZOREK Č. 4 - SÁDROVÝ MODEL BUNKRU

Ze vstupních snímků (Obr. 4.25) objektu a aplikováním kroků zmíněných v předchozí kapitole bylo dosaženo výstupu, který je zaznamenán na Obr. 4.26



Obr. 4.25: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°



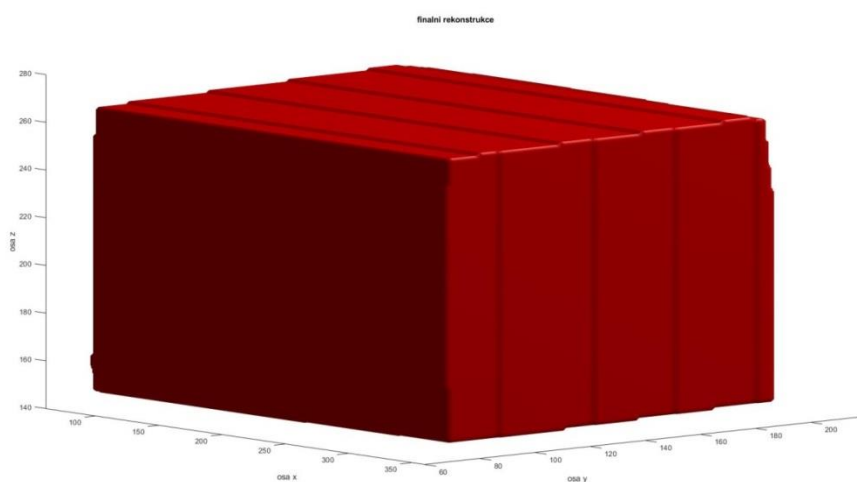
Obr. 4.26: Konečná rekonstrukce modelu bunkru

4.2.6 VZOREK Č. 5 - KRABIČKA OD ČAJE

Za vstupních snímků (Obr. 4.27) objektu a aplikováním kroků zmíněných v předchozí kapitole bylo dosaženo výstupu, který je zaznamenán na Obr. 4.28



Obr. 4.27: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0°, (b) 90°, (c) 270° a (d) o 345°

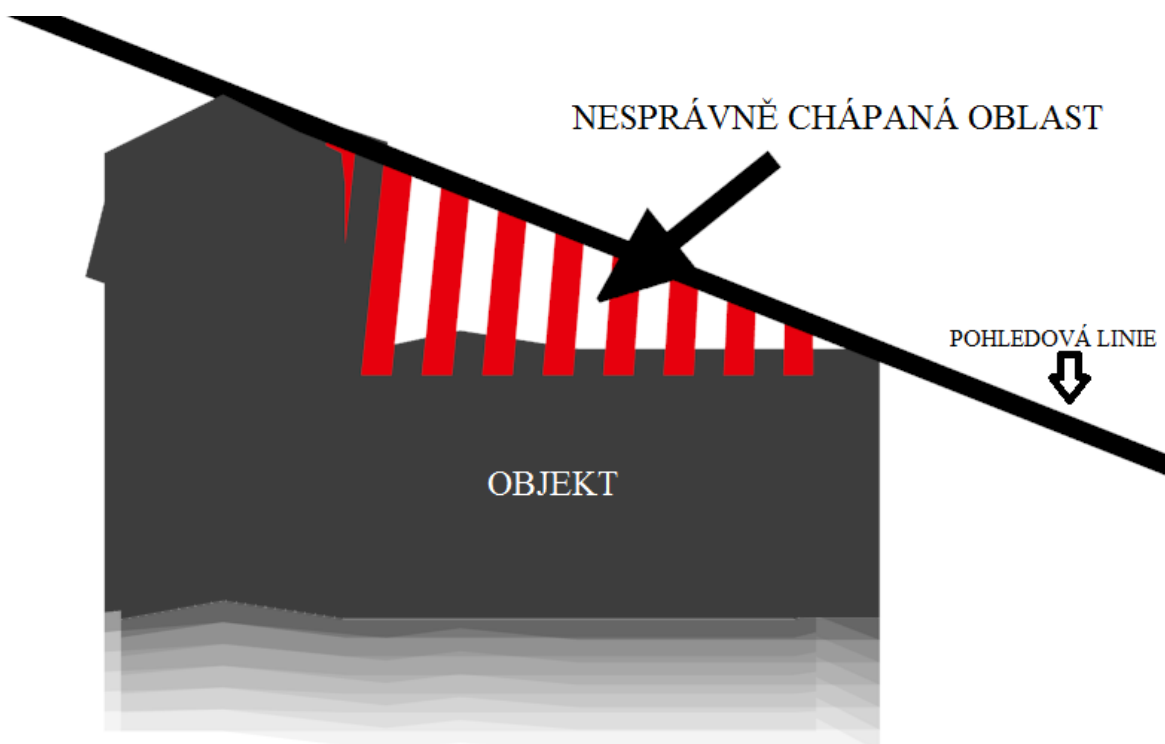


Obr. 4.28: Konečná rekonstrukce krabičky čaje

5 DISKUSE

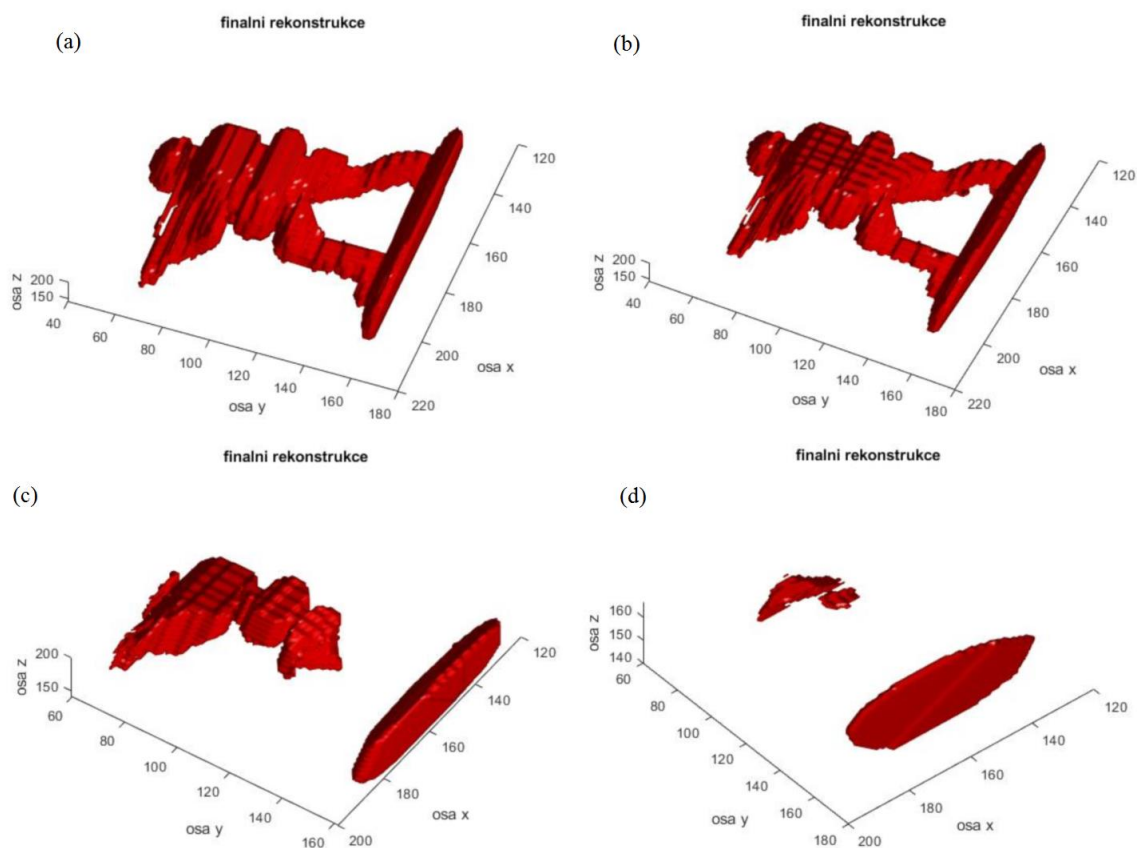
Ze získaných výsledků je patrné, že metoda „shape from silhouettes“ (SfS) dosáhla uspokojivějších výsledků, než SfM. Možným důvodem neúspěchu mohla být málo pestrá textura objektů a nerovnoměrné osvětlení. V případě modelu bunkru, metoda SfM zřejmě selhala z důvodu texturově chudého modelu. Bylo tak detekováno málo bodů, což zkomplikovalo realizaci (Obr. 4.13). U figurky vojáčka byl zřejmě problém mírně odlišný. Ačkoliv byl texturově bohatý (maskování uniformy), byl velmi malý. S kombinací digitalizace snímku fotoaparátem došlo ke splnutí některých částí textury, ztrátě informace. Počítač byl tak schopen detekovat pouze malé množství bodů a byly tak zachyceny pouze okraje (Obr. 4.16). Jediným objektem, který se podařilo rekonstruovat, byla krabička od čaje (Obr. 4.11). Nicméně i zde nebylo konečné mračno bodů příliš podobné rekonstruovanému objektu. Důvodem mohlo být špatné pochopení algoritmu registrace mračna bodů do již registrovaných mračen. Došlo tak k paprskovým artefaktům a přítomnosti přebytečných bodů mimo objekt (čímž se rovněž zhoršila zřetelnost výsledného mračna bodů). Dalším možným důvodem mohla být nesprávná transformace. Ačkoliv se každou iteraci mračno bodů otáčelo po ose-z o známý úhel (15 stupňů), nebyl výstup správný. Zřejmě mohlo dojít k nepřesnému natáčení rekonstruovaného objektu během jeho focení.

Ačkoliv byla metoda SfS úspěšnější, měl výstup jisté nedostatky. Hlavním problémem byla ztráta informace o reliéfu a textuře objektu, z důvodu použití binárních snímků. Ztrátou informace o reliéfu objektu došlo také k tomu, že do rekonstrukce byly zařazené i oblasti, které se nevztahovaly k objektu (Obr. 5.1).

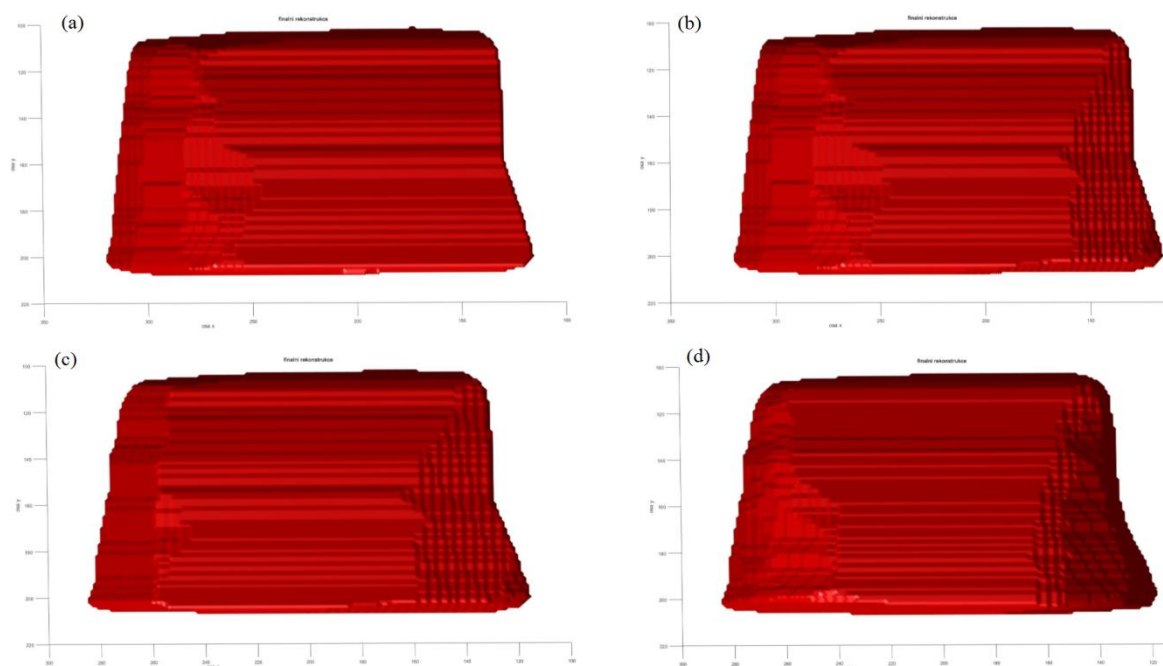


Obr. 5.1: Ilustrativní ukázka špatné rekonstrukce modelu bunkru vlivem nesprávného pochopení struktury objektu vlivem chybějící informace v binárním snímku siluety

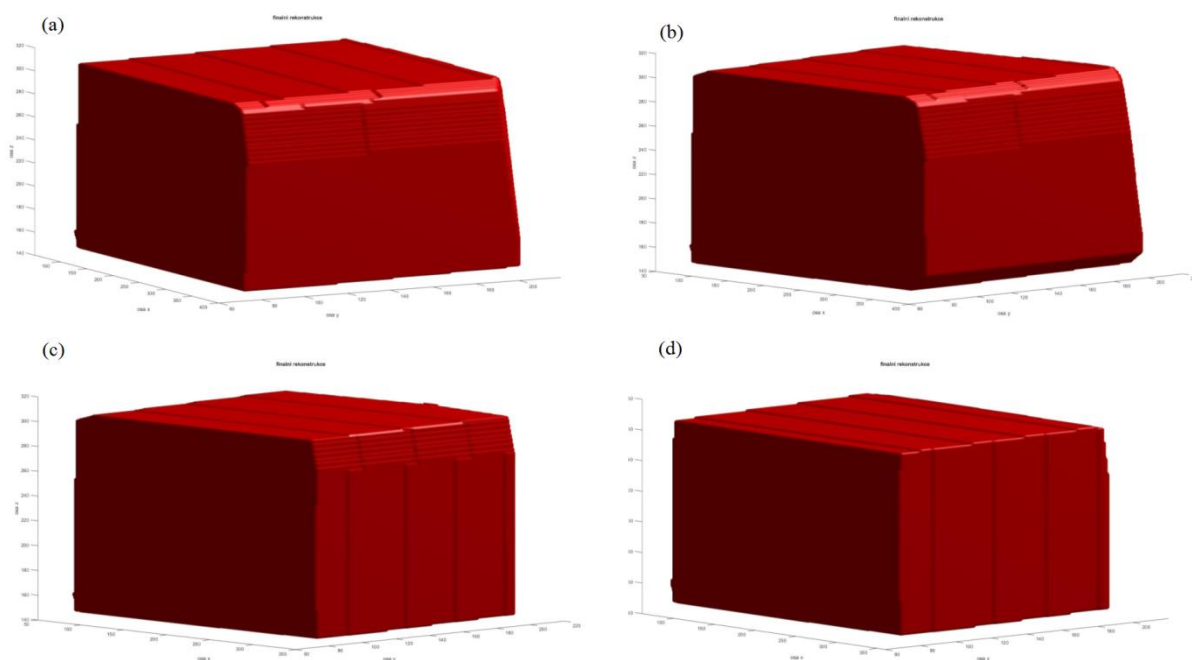
Dalším problémem, který bylo třeba brát na vědomí, bylo zachování stejné vzdálenosti objektu od kamery, tím pádem i přesné umístění objektu ve středu otočného stolku. I posun o několik milimetrů zde hrál roli na úspěšnost konečné rekonstrukce. Ze zmíněného důvodu mohlo dojít k nesprávnému nabývání hodnot u společné oblasti, čímž, následným prahováním, mohlo dojít k odstranění částí objektu, které měly být zachovány. Nejvýraznější problém zmíněného důvodu je u figurky vojáčka, kde pro získání uspokojivého výsledku (Obr. 5.2 (b)) bylo potřeba snížit počet 2D snímků, účastnících se rekonstrukce, z 25-ti na prvních 8 snímků. S přibývajícimi snímky docházelo po následném prahování ke ztrátě společných částí. Nejdříve se odstranily nohy (Obr. 5.2 (c)) a část hlavy, nakonec i spodní polovina trupu (Obr. 5.2 (d)). Naopak s menším počtem snímků byla nižší kvalita výstupu (Obr. 5.2 (a)). V případě modelu bunkru (Obr. 5.3) a krabičky čaje (Obr. 5.4) je vliv správného umístění objektu během snímání menší. Nutno zmínit, že v případě modelu bunkru se do rekonstruovaného objektu přidali části, které s objektem nesouvisí.



Obr. 5.2: Vliv nesprávného umístění objektu během snímání na prahování výsledné rekonstrukce figurky vojáčka : (a) rekonstrukce objektu do 90° (6 snímků), (b) rekonstrukce do 120° (8 snímků), (c) rekonstrukce do 195° (13 snímků), (d) rekonstrukce do 360° (25 snímků)

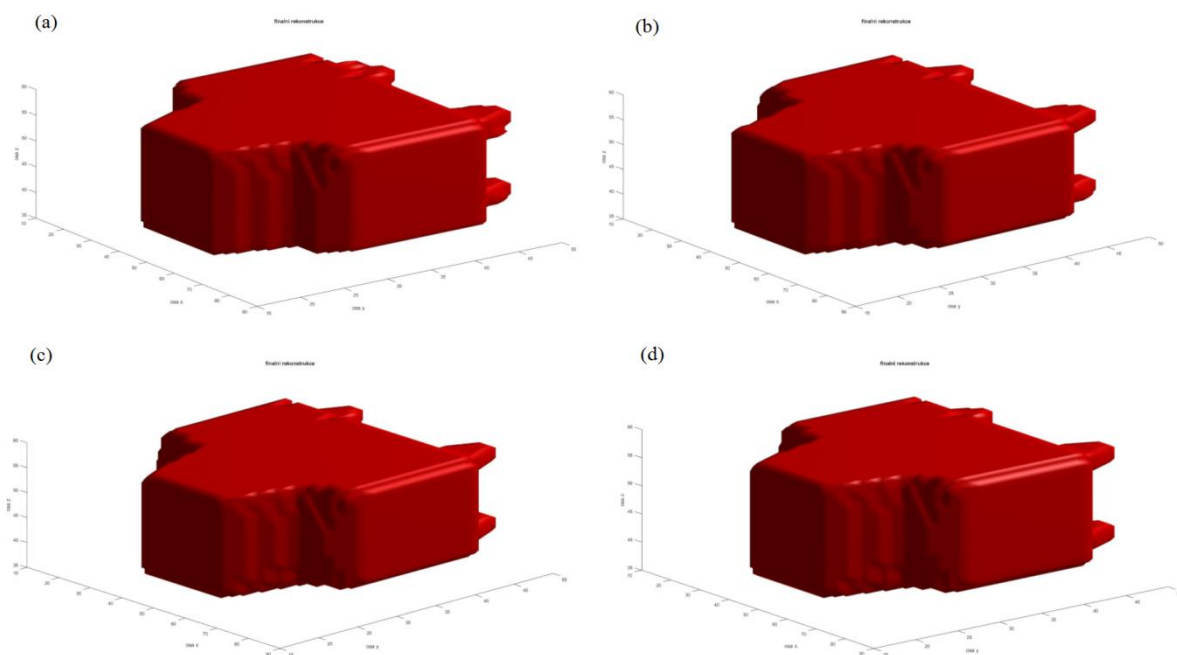


Obr. 5.3: Vliv nesprávného umístění objektu během snímání na prahování výsledné rekonstrukce modelu bunkru : (a) rekonstrukce objektu do 90° (6 snímků), (b) rekonstrukce do 120° (8 snímků), (c) rekonstrukce do 195° (13 snímků), (d) rekonstrukce do 360° (25 snímků)

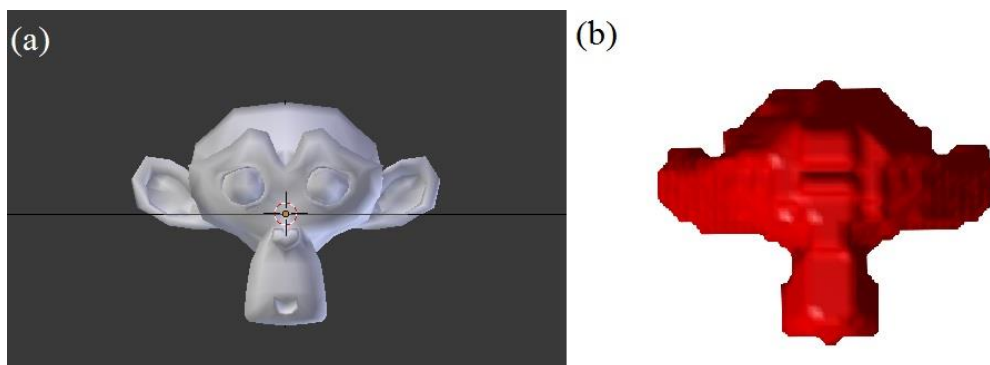


Obr. 5.4: Vliv nesprávného umístění objektu během snímání na prahování výsledné rekonstrukce krabičky čaje : (a) rekonstrukce objektu do 90° (6 snímků), (b) rekonstrukce do 120° (8 snímků), (c) rekonstrukce do 195° (13 snímků), (d) rekonstrukce do 360° (25 snímků)

Figure 1 consists of four subplots labeled (a), (b), (c), and (d), each showing a 3D surface plot of the normalized residual functional connectivity (R-FC) for the left and right hemispheres. The plots are arranged in a 2x2 grid. Each plot has three axes: X (mm) ranging from 0 to 100, Y (mm) ranging from 0 to 100, and Z (mm) ranging from 0 to 100. The surfaces are colored red and show a complex, multi-lobed structure with varying heights, representing the R-FC values across the brain volume.

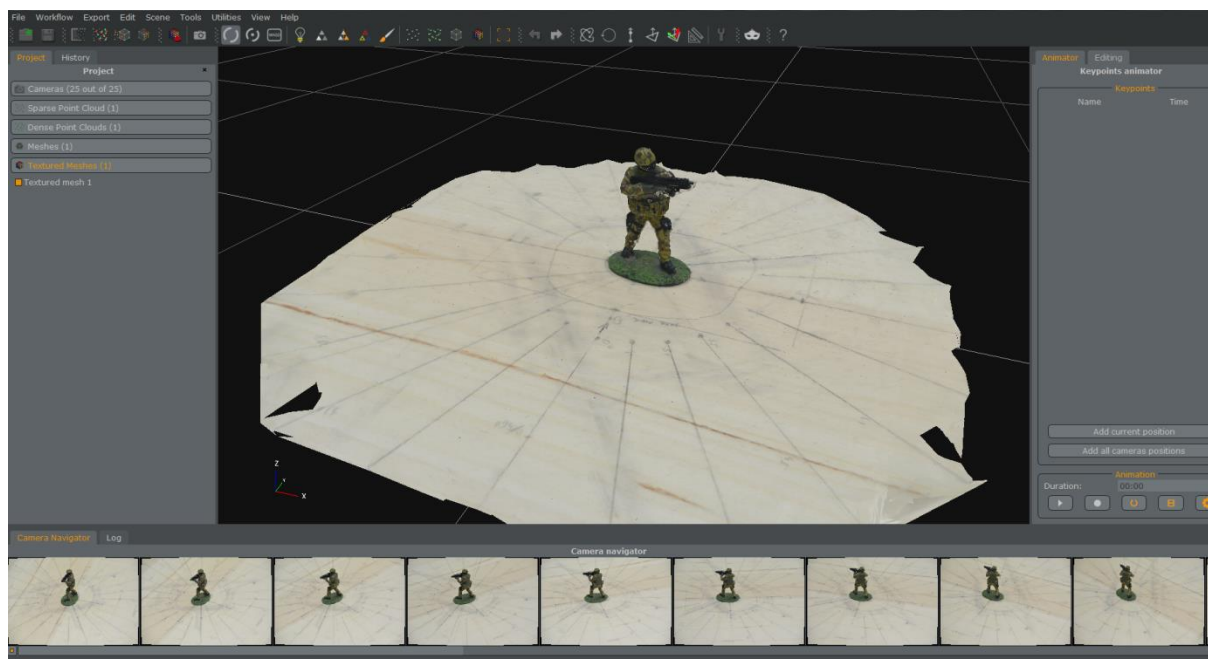


Obr. 5.6: Vliv počtu snímků na konečnou rekonstrukci, virtuální model auta: (a) rekonstrukce objektu do 90° (91 snímků), (b) rekonstrukce do 120° (121 snímků), (c) rekonstrukce do 195° (196 snímků), (d) rekonstrukce do 360° (361 snímků)

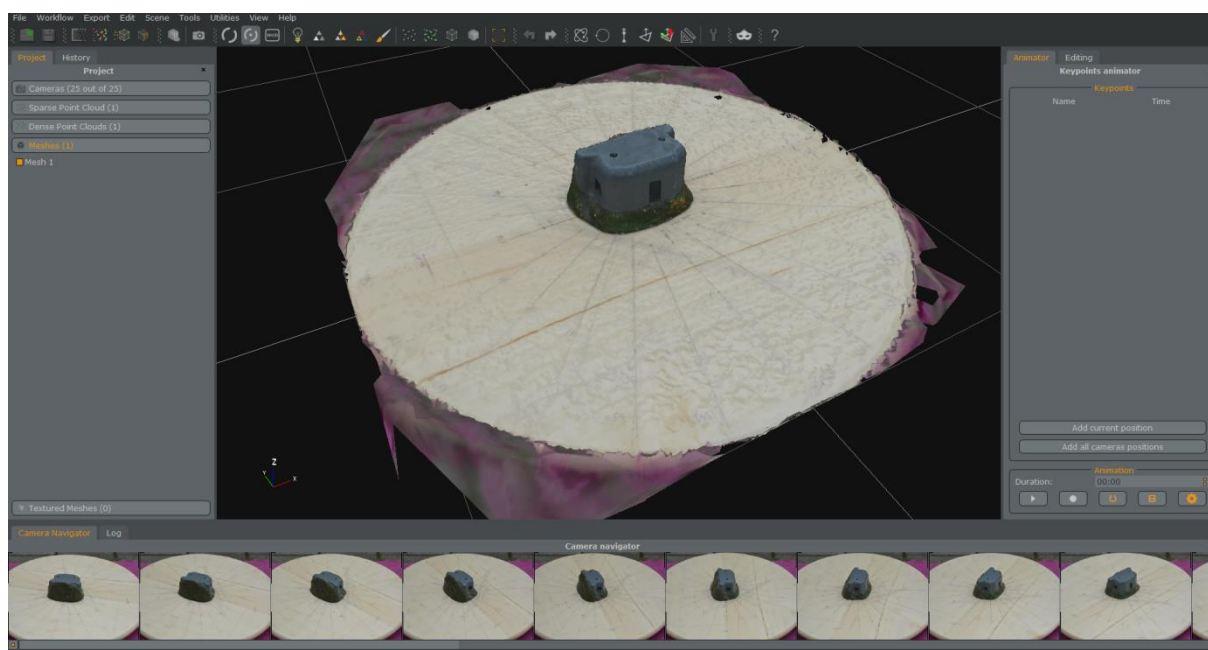


Obr. 5.7: Ztráta informace o reliéfu, textuře a detailů způsobené využitím siluet. (a) původní virtuální model opičky v programu Blender, (b) rekonstruovaný objekt v prostředí MATLAB

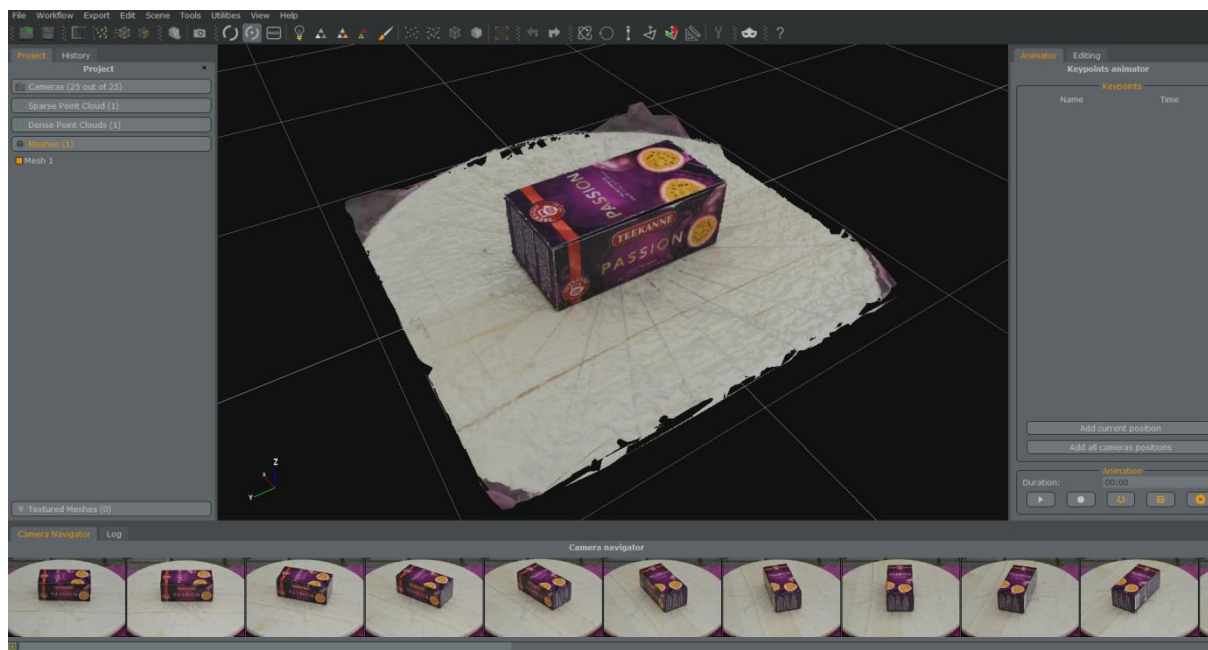
Pro porovnání byly nakonec reálné objekty rekonstruovány komerčním programem 3DF Zephyr verze 4.353, nejen že je schopen provést auto-kalibraci ze vstupních snímků, ale rovněž využívá složitějšího rekonstrukčního přístupu. Nejdříve využívá algoritmu SfM pro výpočet orientace kamer v prostoru, následně se z této informace vypočítá objemové mračno bodů (vícepohledová stereorekonstrukce – „multiple view stereo“). Nakonec z tohoto mračna bodů vypočítá povrch modelu. Z vytvořených rekonstrukcí (Obr 5.8, Obr. 5.9, Obr. 5.10) si lze povšimnout, že program, nejen, že si poradil s auto-kalibrací, ale rovněž rekonstruoval souvislé modely i s texturovou informací pouze s drobnými chybami. Nutno zmínit, že se již jedná o profesionální software tvořený týmem odborníků.



Obr. 5.8: 3D objekt figurky vojáčka rekonstruovaný prostřednictvím programu 3DF Zephyr



Obr. 5.9 : 3D objekt modelu bunkru rekonstruovaný prostřednictvím programu 3DF Zephyr



Obr. 5.10 : 3D objekt krabičky čaje rekonstruovaný prostřednictvím programu 3DF Zephyr

ZÁVĚR

Cílem bakalářské práce bylo seznámení s problematikou 3D rekonstrukce a realizace experimentu na vybraných metodách. V praktické části se zvolila metoda „structure from motion“ (dále jen SfM) a „shape from silhouettes“ (dále jen SfS). Obě metody mají své pro a proti. Jelikož je metoda SfM založená na textuře objektu, bylo nutné, aby byl objekt jednak texturově pestrý a rovněž, aby bylo zachováno jednotné nasvícení objektu ze všech stran, nejlépe bez stínování. Metoda SfS texturově založená není, tím pádem není závislá na pestrosti a nasvícení objektu. Nevýhodou je nicméně ztráta veškerých detailů, zejména problematický byl reliéf a texturová informace objektu, která se v binárních snímcích ztratila.

Dle výsledků a diskuze bylo zjištěno, že metoda SfS je úspěšnější, než metoda SfM. Důvodem mohl být vedle již zmíněného špatného nasvícení a menší barevnosti povrchu objektu také lidský faktor. Využitím metody SfS se podařilo rekonstruovat všechny objekty, kde v případě virtuálních objektů proběhla rekonstrukce bez komplikací. Bylo nicméně zjištěno, že tato metoda je velmi citlivá na umístění objektu vůči kameře, tudíž bylo nutné u objektu figurky vojáčka, pro uspokojivý výsledek, snížit počet rekonstruovaných snímků z 25-ti snímků na prvních 8, jelikož se následující snímky nesprávně sumovaly do rekonstrukční matice. V případě SfM se podařilo rekonstruovat pouze jeden objekt, krabičku od čaje, která ve srovnání s metodou SfS byla nepřesná.

Nakonec se tyto objekty pro porovnání rekonstruovaly v profesionálním programu 3DF Zephyr, ve kterém se objekty nejen rekonstruovaly bez potíží, ale také bez potřeby kalibrace. Bylo rovněž diskutováno možné využití 3D rekonstrukce v medicíně (v úvodu), kde bylo zjištěno, že tato problematika je velmi rozmanitá a nabývá stále většího významu.

LITERATURA

- [1] ŽÁRA, J. *Moderní počítačová grafika*. 2., přeprac. a rozš. vyd. Brno: Computer Press, 2004. ISBN 80-251-0454-0.
- [2] SZELISKI, R. *Computer Vision: Algorithms and Applications*. London: Springer, 2010. Texts in computer science. ISBN 1848829353.
- [3] ZHANG, Z. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 2000, 22.
- [4] GUERCHOUCHE, R., COLDEFY, F. Camera Calibration Methods Evaluation Procedure for Images Rectification and 3D Reconstruction. In : *WSCG ...: the ... International conference in Central Europe on computer graphics and visualization ... : conference proceedings*. Plzeň: University of West Bohemia. ISBN 978-80-86943-16-9.
- [5] HEYDEN, A.; POLLEFEYS, M. Multiple view geometry. *Emerging topics in computer vision*, 2005, 45-107.
- [6] MOONS, T., et al. 3D reconstruction from multiple images part 1: Principles. *Foundations and trends® in Computer Graphics and Vision*, 2010, 4(4), 287-404.
- [7] BIANCO, G., et al. A comparative analysis between active and passive techniques for underwater 3D reconstruction of close-range objects. *Sensors*, 2013, 13(8), 11007-11031.
- [8] WALEK, P.; LAMOŠ, M.; JAN, J. Analýza biomedicínských obrazů. *Analýza biomedicínských obrazů*. Brno, 2013, 1-138.
- [9] HORÁK, K.; KALOVÁ, I. Počítačové vidění.[počítačová cvičení]. *Skripta VUT Brno*. Brno, 2006.
- [10] JURÁNEK, R. *Extrakce obrazových příznaků* [online]. Brno: Vysoké učení technické v Brně, Fakulta informačních technologií. Ústav počítačové grafiky a multimedií, 2010- [cit. 25.12.2018]. Dostupné z: http://www.fit.vutbr.cz/study/courses/IKR/public/stare_prednasky_2011/04_obrazove_priznaky_2010/2010-IKR-ImageFeatures.pdf
- [11] Computer Vision Toolbox. *Computer Vision Toolbox Documentation* [online]. [cit. 25.12.2018]. Dostupné z: <https://www.mathworks.com/help/vision/index.html>

- [12] Camera Calibrator. *Single Camera Calibrator App - MATLAB & Simulink* [online]. [cit. 25.12.2018]. Dostupné z: <https://www.mathworks.com/help/vision/ug/single-camera-calibrator-app.html>
- [13] Camera Calibrator. *What Is Camera Calibration? - MATLAB & Simulink* [online]. [cit. 26.12.2018]. Dostupné z: <https://www.mathworks.com/help/vision/ug/camera-calibration.html>
- [14] BOUGUET, J. Y. *Camera Calibration Toolbox for Matlab* [online]. [cit. 26.12.2018]. Dostupné z: http://www.vision.caltech.edu/bouguetj/calib_doc/
- [15] ZAPRYANOV, G.; NIKOLOVA, I. Comparison and evaluation of edge and corner detectors. Sofia. 2007
- [16] WU, C., et al. High-quality shape from multi-view stereo and shading under general illumination. In: *CVPR 2011*. IEEE, 2011. s. 969-976.
- [17] HARRIS, C. G., et al. A combined corner and edge detector. In: *Alvey vision conference*. 1988. s. 10-5244.
- [18] DEY, N., et al. A comparative study between Moravec and Harris corner detection of noisy images using adaptive wavelet thresholding technique. *arXiv preprint arXiv:1209.1558*, 2012.
- [19] HARTLEY, R., ZISSERMAN A. *Multiple view geometry in computer vision*. 2nd ed. New York: Cambridge University Press, 2003. ISBN 978-0-521-54051-3.
- [20] CHIU, Sh.; CHIU, Chung-Cheng; XU, Sendren. A Background Subtraction Algorithm in Complex Environments Based on Category Entropy Analysis. *Applied Sciences*, 2018, **8**(6), 885.
- [21] AANÆS, H.. Methods for structure from motion. *IMM, Informatik og Matematisk Modelling, Danmarks Tekniske Universitet*, 2003.
- [22] TAPPA, K.; JAMMALAMADAKA, Udayabhanu. Novel biomaterials used in medical 3D printing techniques. *Journal of functional biomaterials*, 2018, **9**(1), 17.
- [23] BAKER, S., et al. Shape-from-silhouette across time part ii: Applications to human modeling and markerless motion tracking. *International Journal of Computer Vision*, 2005, **63**(3), 225-245.

- [24] Ullmann's Encyclopedia of Industrial Chemistry [online]. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co., 2000 [cit. 2019-04-23]. ISBN 3527306730.
- [25] LEUTENEGGER, S.; CHLI, M.; SIEGWART, R. BRISK: Binary robust invariant scalable keypoints. In: *2011 IEEE international conference on computer vision (ICCV)*. Ieee, 2011. s. 2548-2555.
- [26] ESTEBAN, C. H.; SCHMITT, F. Silhouette and stereo fusion for 3D object modeling. *Computer Vision and Image Understanding*, 2004, **96**(3), 367-392.
- [27] CIPOLLA, R.; BLAKE, A. Surface shape from the deformation of apparent contours. *International journal of computer vision*, 1992, **9**(2), 83-112.
- [28] MULAYIM, A. Y.; YILMAZ, U.; ATALAY, V. Silhouette-based 3-D model reconstruction from multiple images. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2003, **33**(4), 582-591.
- [29] ROTH, G., 2015. *Camera Calibration*. [online] Release Date: Aug 26, [cit. 1.4.2019] Dostupné z: http://people.scs.carleton.ca/~c_shu/Courses/comp4900d/notes/camera_calibration.pdf, 2006
- [30] JI, Y. F.; CHANG, C. C. Vision-based sensing for assessing and monitoring civil infrastructures. In: *Sensor Technologies for Civil Infrastructures*. Woodhead Publishing, 2014. s. 383-409.

SEZNAM OBRÁZKŮ

Obr. 1.1: Perspektivní projekce na dírkovém modelu kamery

Obr. 1.2: Ortografická projekce kamery

Obr. 1.3: Radiální zkreslení

Obr. 1.4: Kalibrace kamery. Využití non-plannar pattern (vlevo) a využití plannar pattern (vpravo)

Obr. 1.5: Původní obrázek (vpravo), ukázka hranové detekce aproximující druhou derivaci pomocí lokálního operátoru Laplacián (vpravo)

Obr. 1.6: Ilustrace řešení problematiky vyhledávání vhodného významného bodu pomocí posuvného okna

Obr. 1.7: Rohová detekce pomocí Harris–Stephensesnova algoritmu s 50-ti detekovanými rohy

Obr. 1.8: Metoda segmentace pomocí zachování popředí

Obr. 1.9: Chromatické klíčování, ve vstupním snímku je veškerá zelená vyjmuta, obraz má následné průhledné pozadí, následným vrstvením s novým snímkem vznikne konečný obraz

Obr. 1.10: Problematika epipolární geometrie ilustrována na dvou pohledech

Obr. 1.11: Základní taxonomie metod pro extrakci 3D informace zkoumaného objektu

Obr. 1.12: Shape-from-silhouette pomocí „voxel carving“

Obr. 2.1: Návrh akvizičního systému

Obr. 2.2: Návrh akvizičního systému v programu Blender

Obr. 3.1: Vygenerovaná šachovnice pro kalibraci kamery

Obr. 3.2: Ukázka prostředí Camera Calibrator App s nahranými snímky kalibrační šachovnice (před spuštěním kalibrace).

Obr. 4.1: Ukázka detekce rohu (zelená kolečka) a jejich reprojekce (červené křížky)

Obr. 4.2: (a) Bližší pohled na výsledky odhadu reprojekčních chyb

Obr. 4.3: Odhadnuté vnější parametry kamery vizualizované v prostředí MATLAB

Obr. 4.3: Ukázka vstupních snímků rekonstruovaného objektu

Obr. 4.4: Převedení barevných snímků na snímky šedotónové

Obr. 4.5: Korekce snímků pomocí příkazu `undistortImage`

Obr. 4.6: Detekované rohy pomocí Harris-Stephensonova algoritmu u prvního snímku.

Obr. 4.7: Hrubé nalezení korespondujících bodů mezi první dvojicí snímků

Obr. 4.8: Vymezení korespondujících bodů u první dvojice snímků

Obr. 4.9: 3D rekonstrukce první dvojice (natočení objektu o 0° a 15°)

Obr. 4.10: 3D rekonstrukce šesté dvojice (natočení objektu o 75° a 90°)

Obr. 4.11: Konečná podoba výsledného point cloudu krabičky čaje

Obr. 4.13: Ukázka detekovaných rohů u prvního snímku

Obr. 4.14 : Neúspěšná rekonstrukce objektu dána vlivem nedostatečného množství nalezených korespondujících bodů

Obr. 4.15: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

Obr. 4.16: Ukázka detekovaných rohů u prvního snímku

Obr. 4.17 : Neúspěšná rekonstrukce objektu dána vlivem nedostatečného množství nalezených korespondujících bodů

Obr. 4.18: Ilustrativní ukázka vytvořené siluety převedením RGB snímku na šedotónový a následným uplatněním prahování.

Obr. 4.19: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

Obr. 4.20: Konečná rekonstrukce virtuálního modelu auta

Obr. 4.21: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

Obr. 4.22: Konečná rekonstrukce virtuálního modelu opičky

Obr. 4.23: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

Obr. 4.24: Konečná rekonstrukce figurky vojáčka

Obr. 4.25: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

Obr. 4.26: Konečná rekonstrukce modelu bunkru

Obr. 4.27: Ukázka vstupních snímků rekonstruovaného objektu, natočeného o (a) 0° , (b) 90° , (c) 270° a (d) o 345°

Obr. 4.28: Konečná rekonstrukce krabičky čaje

Obr. 5.1: Ilustrativní ukázka špatné rekonstrukce modelu bunkru vlivem nesprávného pochopení struktury objektu vlivem chybějící informace v binárním snímku siluety

Obr. 5.2: Vliv nesprávného umístění objektu během snímání na prahování výsledné rekonstrukce figurky vojáčka

Obr. 5.3: Vliv nesprávného umístění objektu během snímání na prahování výsledné rekonstrukce modelu bunkru

Obr. 5.4: Vliv nesprávného umístění objektu během snímání na prahování výsledné rekonstrukce krabičky čaje

Obr. 5.6: Vliv počtu snímků na konečnou rekonstrukci, virtuální model auta

Obr. 5.7: Ztráta informace o reliéfu, textuře a detailů způsobené využitím siluet

Obr. 5.8: 3D objekt figurky vojáčka rekonstruovaný prostřednictvím programu 3DF Zephyr

Obr. 5.9 : 3D objekt modelu bunkru rekonstruovaný prostřednictvím programu 3DF Zephyr

Obr. 5.10 : 3D objekt krabičky čaje rekonstruovaný prostřednictvím programu 3DF Zephyr

OBSAH CD

Obsahem přiloženého CD je bakalářská práce ve formátu .PDF a dva .ZIP soubory. Structure_from_motion.ZIP obsahuje realizované pokusy 3D rekonstrukce použitím metody „structure from motion“. Obsahuje .m soubor pro převzorkování snímků a čtyři složky. Složka calibracni_obj obsahuje snímky kalibrační matice a .m soubor, který tento kalibrační objekt vytvořil. Zbýlé tři složky obsahují aplikované pokusy 3D rekonstrukce jednotlivých objektů (vstupní snímky, .m soubor, parametry kamery). Soubor Shape_from_silhouettes.ZIP rovněž obsahuje .m soubor pro převzorkování snímků, nicméně je zde složek pět. Všechny složky obsahují realizované pokusy 3D rekonstrukce pomocí metody „shape from silhouettes“ (vstupní snímky a .m soubor), kde dvě složky rekonstruují virtuální objekty a zbylé tři objekty skutečné. Z důvodu vysoké velikosti jednotlivých snímků se v CD nachází již snímky podvzorkované. Součástí každého .ZIP souboru je rovněž textový dokument README.txt, který obsahuje stručný obsah daného souboru.

Soubory:

RehorBP.pdf

Structure_from_motion.zip

Shape_from_silhouettes.zip